HERIOT WATT UNIVERSITY

Greedy Techniques for Magnetic Resonance Fingerprinting

Author:
Roberto de Jesús Duarte Coello

Supervisor:
Prof. Yves Wiaux

A thesis submitted for the degree of Doctor of Philosophy

Heriot-Watt University
School of Engineering and Physical Sciences
Institute of Sensors, Signals and Systems

May 2020

Abstract

In this manuscript, we show four main results in the context of Magnetic Resonance Fingerprinting (MRF):

- A memory efficient method to explore the manifold of fingerprints.
- A method that allows super-resolution reconstructions relying on spatial regularisation.
- An extension to partial volumes and a greedy approximate projection algorithm.
- An extension to Self-Calibration and Imaging.

In quantitative Magnetic Resonance Imaging, traditional methods suffer from the so-called Partial Volume Effect (PVE) due to spatial resolution limitations. As a consequence of PVE, the parameters of the voxels containing more than one tissue are not correctly estimated. MRF is not an exception. The existing methods addressing PVE are neither scalable nor accurate. We propose to formulate the recovery of multiple tissues per voxel as a non-convex constrained least-squares minimisation problem. To solve this problem, we develop a memory efficient, greedy approximate projected gradient descent algorithm, dubbed GAP-MRF. Our method adaptively finds the regions of interest on the manifold of fingerprints defined by the MRF sequence. We generalise our method to compensate for phase errors appearing in the model, using an alternating minimisation approach. We show, through simulations on synthetic data with PVE, that our algorithm outperforms state-of-the-art methods in reconstruction quality. Our approach is validated on the EUROSPIN phantom and on *in vivo* datasets.

Coil sensitivity calibration is a crucial step in the reconstruction process to obtain accurate results. Usual MRI self-calibration methods, reconstructing independently the time acquisitions, are not suitable for highly undersampled MRF data. In this work, leveraging recent developments in non-convex optimisation, we propose the first self-calibration method for MRF, exploiting the correlation in the time acquisitions,

the spatial regularity of the magnetisation images and the smoothness of the coil sensitivity maps.

Keywords: Magnetic Resonance Imaging, MR Fingerprinting, Inverse Problems, Self-Calibration and Imaging, Partial Volumes, Alternating Minimisation.

Research Thesis Submission

Please note this form should be bound into the submitted thesis.

Name:	ROBERTO DE JESUS DUARTE COELLO		
School:	EPS		
Version: *(i.e. First, Resubmission, Final)*	FINAL SUBMISSION	Degree Sought:	DOCTOR OF PHILOSOPHY

Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

1. The thesis embodies the results of my own work and has been composed by myself
2. Where appropriate, I have made acknowledgement of the work of others
3. The thesis is the correct version for submission and is the same version as any electronic versions submitted*.
4. My thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying and be available via the Institutional Repository, subject to such conditions as the Librarian may require
5. I understand that as a student of the University I am required to abide by the Regulations of the University and to conform to its discipline.
6. I confirm that the thesis has been verified against plagiarism via an approved plagiarism detection application e.g. Turnitin.

ONLY for submissions including published works

Please note you are only required to complete the Inclusion of Published Works Form (page 2) if your thesis contains published works)

7. Where the thesis contains published outputs under Regulation 6 (9.1.2) or Regulation 43 (9) these are accompanied by a critical review which accurately describes my contribution to the research and, for multi-author outputs, a signed declaration indicating the contribution of each author (complete)
8. Inclusion of published outputs under Regulation 6 (9.1.2) or Regulation 43 (9) shall not constitute plagiarism.

* *Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.*

Signature of Candidate:		Date:	06/05/2020

Submission

Submitted By *(name in capitals):*	
Signature of Individual Submitting:	
Date Submitted:	

For Completion in the Student Service Centre (SSC)

Limited Access	Requested	Yes		No		Approved	Yes		No	
*E-thesis Submitted (**mandatory for final theses**)*										
Received in the SSC by *(name in capitals):*						Date:				

Inclusion of Published Works

Please note you are only required to complete the Inclusion of Published Works Form if your thesis contains published works under Regulation 6 (9.1.2)

Declaration

This thesis contains one or more multi-author published works. In accordance with Regulation 6 (9.1.2) I hereby declare that the contributions of each author to these publications is as follows:

Citation details	Duarte Coello, R. D. J., Chen, Z., Golbabaee, M., Mahbub, Z., Marshall, I., Davies, M., & Wiaux, Y. (2017). Adaptive-BLIP for Magnetic Resonance Fingerprinting. Paper presented at International Biomedical and Astronomical Signal Processing Frontiers Workshop 2017, Villars-sur-Ollon, Switzerland.
Author 1	Conception and design of the work, coding, analysis and interpretation of data for the work.
Author 2	Collaboration and writing assistance, technical editing, language editing, and proofreading
Author 3	Scientific discussion and proofreading
Author 4	Scientific discussion and proofreading
Author 5	Scientific discussion and proofreading
Author 6	Scientific discussion and proofreading
Author 7	Problem definition, general supervision, language editing, and proofreading
Signature:	
Date:	06/05/2020

Citation details	Duarte Coello, R. D. J., Chen, Z., Gazzola, S., Marshall, I., Davies, M., & Wiaux, Y. (2017). Convex optimisation for partial volume estimation in compressive quantitative MRI. In 6th Signal Processing with Adaptive Sparse Structured Representations workshop 2017
Author 1	Conception and design of the work, coding, analysis and interpretation of data for the work.
Author 2	Collaboration and writing assistance, technical editing, language editing, and proofreading
Author 3	Collaboration and writing assistance
Author 4	Scientific discussion and proofreading
Author 5	Scientific discussion and proofreading
Author 6	Problem definition, general supervision, language editing, and proofreading
Signature:	
Date:	06/05/2020

Citation details	Duarte Coello, R. D. J., Chen, Z., Davies, M. E., & Wiaux, Y. (2017). Adaptive-BLIP for partial volume reconstructions in Magnetic Resonance Fingerprinting. Paper presented at ISMRM Workshop on Magnetic Resonance Fingerprinting, Cleveland, United States.
Author 1	Conception and design of the work, coding, analysis and interpretation of data for the work.
Author 2	Collaboration and writing assistance, technical editing, language editing, and proofreading
Author 3	Scientific discussion and proofreading
Author 4	Problem definition, general supervision, language editing, and proofreading
Signature:	
Date:	06/05/2020

Citation details	Duarte Coello, R. D. J., Repetti, A., & Wiaux, Y. (2019). Self-calibration for Magnetic Resonance Fingerprinting. Paper presented at International BASP Frontiers workshop 2019, Villars sur Ollon, Switzerland.
Author 1	Conception and design of the work, coding, analysis and interpretation of data for the work.
Author 2	Collaboration and writing assistance, technical editing, language editing, and proofreading
Author 3	Problem definition, general supervision, language editing, and proofreading
Signature:	
Date:	06/05/2020

Citation details	Duarte, R., Repetti, A., Gómez, P. A., Davies, M., & Wiaux, Y. (2020). Greedy approximate projection for magnetic resonance fingerprinting with partial volumes. Inverse Problems, 36(3), 035015.
Author 1	Conception and design of the work, coding, analysis and interpretation of data for the work.
Author 2	Collaboration and writing assistance, technical editing, language editing, and proofreading
Author 3	Data acquisition, scientific discussion and proofreading
Author 4	Scientific discussion and proofreading
Author 5	Problem definition, general supervision, language editing, and proofreading
Signature:	
Date:	06/05/2020

Please included additional citations as required.

Contents

List of Figures

List of Tables

N	Number of voxels in the volume of interest
Q	Number of measurements per excitation and per coil
C	Number of coils
L	Number of excitation instances
$\mathbf{Y} \in \mathbb{C}^{Q \times L \times C}$	Measurement matrix
$\mathbf{M} \in \mathbb{C}^{N \times L}$	Magnetisation Response of the volume of interest, introduced in (3.1)
h	Linear operator from $\mathbb{C}^{N \times L}$ to $\mathbb{C}^{Q \times L \times C}$ defining the acquisition process
P	Number of the tissue parameters
$\mathcal{M} \subset \mathbb{R}^{1 \times P}$	Subset describing the feasible parameter space
B	Non-linear smooth operator from $\mathcal{M}$ to $\mathbb{C}^{1 \times L}$ describing the magnetic resonance experiment
$\mathbf{\Phi} \in \mathbb{C}^{D \times L}$	Discretisation of B with D elements
$\mathcal{B}^+$	Set describing $\mathbf{M}$ as a magnetisation response of pure voxels
$\mathbf{X} \in \mathbb{R}_+^{N \times D}$	Mixing matrix used to describe the PVE
$\mathcal{B}_{\mathcal{S}_+}(\mathbf{\Phi})$	Set describing all possible $\mathbf{M}$ satisfying the proposed PV model, defined in (4.17)
G	Magnetisation sequence spatial regularisation function for the rABLIP algorithm
$\mathcal{S}_+$	Set describing the intersection of the sets $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$ and $\mathcal{S}_4$, defined in (4.18)
$\mathcal{S}_1$	Positive orthant, defined in (4.19)
$\mathcal{S}_2$	Set describing the maximum number of active dictionary elements, defined in (4.20) with: $\mathcal{G}_{\mathbf{X}}$ Set indicating the columns of $\mathbf{X}$ contributing to the magnetisation response ξ Minimum voxel proton density $\mathcal{D}_{\mathbf{X}}$ Set giving the voxels with significant contribution to the magnetisation response K Maximum number of active dictionary elements
$\mathcal{S}_3$	Set describing the constraint on the distinct active dictionary elements, defined in (4.21) with: υ Constant used to define neighbourhoods for the dictionary elements in the parameter space $\mathcal{N}_\upsilon$ Set describing the neighbour dictionary elements for given dictionary element d
$\mathcal{S}_4$	Set giving the minimum number of pure voxels per active dictionary element, defined in (4.23) with: $\mathcal{V}_{\mathbf{X}}$ Set describing the pure voxels in $\mathbf{X}$ κ Minimum number of pure voxels per active dictionary element
μ	Gradient step size
$\mathbf{U} \in \mathbb{R}^{N \times \overline{T}}$	Sub-matrix of $\mathbf{X}$
$\mathbf{\Delta} \in \mathbb{C}^{\overline{T} \times L}$	Sub-matrix of $\mathbf{\Phi}$
$\mathcal{Z}$	Operator from $\mathbb{R}_+^{N \times \overline{T}}$ to $\mathbb{R}_+^{N \times D}$ that maps a matrix $\mathbf{U}$ to the corresponding $\mathbf{X}$
$\boldsymbol{\lambda} \in \mathbb{C}^N$	Vector used to compensate for the complex phase errors in the model
γ	Tolerance parameter for a pure voxels
$\tau, \tau_k, \tau_\upsilon, \tau_\kappa$	Tolerance parameters for the initialisation process
G_{TV}	Magnetisation sequence regularisation function for the Self-Calibrated GAP-MRF
$G_{\mathbf{S}}$	Coil sensitivity map regularisation function for the Self-Calibrated GAP-MRF

Chapter 1

Introduction

Advances in technology and the availability of relatively inexpensive computational resources have allowed combining data acquisition and image processing. This set of techniques are called computational imaging. Computational imaging has acquired an important role in diagnosis in the last decades [1]. These techniques are preferred by clinicians because they provide important diagnostic information while been less-invasive or non-invasive to the patient. This manuscript focuses on a modality of the Magnetic Resonance Imaging (MRI) technique called quantitative MRI (qMRI). qMRI aims to retrieve the parameters that characterise the imaged volume of interest in a feasible time. In particular, this manuscript target the inverse problem stated by the recently proposed acquisition procedure called Magnetic Resonance Fingerprinting (MRF) in [2]. The objective of this manuscript is to analyse in detail the MRF inverse problem and propose an efficient technique that addresses the main challenges associated with this acquisition. In this thesis, we aim to solve non-convex minimisation problems with the following structure:

$$\underset{\substack{\mathbf{M}\in\mathbb{C}^{N\times L}\\ \mathbf{S}\in\mathbb{C}^{N\times C}}}{\text{minimise}} \underbrace{F\left(\mathbf{S},\mathbf{M}\right)}_{\text{data fidelity}} + \underbrace{\iota_{\mathcal{T}}\left(\mathbf{M}\right)}_{\substack{\text{magnetisation}\\ \text{temporal}\\ \text{regularisation}}} + r_1 \underbrace{G_{\mathbf{M}}\left(\mathbf{M}\right)}_{\substack{\text{magnetisation}\\ \text{spatial}\\ \text{regularisation}}} + r_2 \underbrace{G_{\mathbf{S}}\left(\mathbf{S}\right)}_{\substack{\text{coil sensitivity}\\ \text{spatial}\\ \text{regularisation}}}. \tag{1.1}$$

This framework not only targets the imaging problem with partial volumes but also the coil sensitivity calibration problem. Based on the recent advances in optimisation, we propose greedy approximate projection algorithms that allow to solve the problem efficiently. The proposed algorithmic structure is based on the algorithm presented in [3] and can be seen in Algorithm 1.

Algorithm 1 Self-Calibrated GAP-MRF

Input: $\mathbf{S}^{(0)} \in \mathbb{C}^{N \times C}$, $\mathbf{M}^{(0)} \in \mathcal{T}$, $(I, J) \in \mathbb{N}_*^2$, $(\mu, \nu) \in (\mathbb{R}_*^+)^2$
for $k = 0, 1, \ldots$ **do**
  $\widetilde{\mathbf{M}}^{(0)} = \mathbf{M}^{(k)}$ and $\widetilde{\mathbf{S}}^{(0)} = \mathbf{S}^{(k)}$
  for $i = 0, 1, \ldots, I-1$ **do**
    $\widetilde{\mathbf{M}}^{(i+1)} \approx \operatorname{prox}_{r_1 \mu^{(i)} G_{\mathbf{M}}(\cdot)} \left(\mathcal{P}_{\mathcal{T}} \left(\widetilde{\mathbf{M}}^{(i)} - \mu^{(i)} \nabla_{\mathbf{M}} F \left(\mathbf{S}^{(k)}, \widetilde{\mathbf{M}}^{(i)} \right) \right) \right)$
  $\mathbf{M}^{(k+1)} = \widetilde{\mathbf{M}}^{(I)}$
  for $j = 0, 1, \ldots, J-1$ **do**
    $\widetilde{\mathbf{S}}^{(j+1)} = \operatorname{prox}_{r_2 \nu G_{\mathbf{S}}(\cdot)} \left(\widetilde{\mathbf{S}}^{(j)} - \nu \nabla_{\mathbf{S}} F \left(\widetilde{\mathbf{S}}^{(j)}, \mathbf{M}^{(k+1)} \right) \right)$
  $\mathbf{S}^{(k+1)} = \widetilde{\mathbf{S}}^{(J)}$

The main results of this manuscript are presented in the following publications:

- Duarte Coello, R. D. J., Chen, Z., Golbabaee, M., Mahbub, Z., Marshall, I., Davies, M., & Wiaux, Y. (2017). Adaptive-BLIP for Magnetic Resonance Fingerprinting. Paper presented at International Biomedical and Astronomical Signal Processing Frontiers Workshop 2017, Villars-sur-Ollon, Switzerland.

- Duarte Coello, R. D. J., Chen, Z., Gazzola, S., Marshall, I., Davies, M., & Wiaux, Y. (2017). Convex optimisation for partial volume estimation in compressive quantitative MRI. In 6th Signal Processing with Adaptive Sparse Structured Representations workshop 2017.

- Duarte Coello, R. D. J., Chen, Z., Davies, M. E., & Wiaux, Y. (2017). Adaptive-BLIP for partial volume reconstructions in Magnetic Resonance Fingerprinting. Paper presented at ISMRM Workshop on Magnetic Resonance Fingerprinting, Cleveland, United States.

- Duarte Coello, R. D. J., Repetti, A., & Wiaux, Y. (2019). Self-calibration for Magnetic Resonance Fingerprinting. Paper presented at International BASP Frontiers workshop 2019, Villars sur Ollon, Switzerland.

- Duarte, R., Repetti, A., Gómez, P. A., Davies, M., & Wiaux, Y. (2020). Greedy approximate projection for magnetic resonance fingerprinting with partial volumes. Inverse Problems, 36(3), 035015.

The manuscript is organised as follows: the first chapter introduces the mathematical notation used in the manuscript, the basics of inverse problems and MRI physics. The second chapter focuses on the optimisation tools used to solve inverse problems. In the third chapter, the MRF technique is described as an inverse problem, the state-of-the-art methods are described and a greedy method is proposed. In the fourth chapter, the Partial Volume Effect (PVE) is incorporated into the model, the state-of-the-art methods are described and the proposed method is extended to partial volumes. In the fifth chapter, we present the need for accurate calibration and we propose an extension of our method to self-calibration and imaging. Finally, the conclusions of this work are presented in the last chapter.

1.1 Notation

In this section, we introduce the notation we will use in the remainder of the paper. We refer the reader to [4, 5] for additional details about optimisation. To have a compact notation when selecting a specific row $n \in \{1, \dots, N\}$ of a matrix $\mathbf{M} \in \mathbb{C}^{N \times L}$, we use the notation $\mathbf{M}_{n,:} = (\mathbf{M}_{n,l})_{1 \leq l \leq L}$. Similarly, to select a specific column $l \in \{1, \dots, L\}$ of this matrix, we use $\mathbf{M}_{:,l} = (\mathbf{M}_{n,l})_{1 \leq n \leq N}$. More generally, this notation is also used to select subparts of tensors. The operator $\text{real}(\cdot)$ gives the real part of its complex argument, the operator $\text{Diag}(\cdot)$ builds a diagonal matrix whose diagonal elements are given by its argument, and $(\cdot)^\dagger$ gives its adjoint. The adjoint of a linear operator $g \colon \mathbb{C}^L \to \mathbb{C}^N$ is denoted by $g^\dagger$. The adjoint of a linear operator satisfies:

$$\langle g(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, g^\dagger(\mathbf{y}) \rangle, \tag{1.2}$$

$\forall \mathbf{x} \in \mathbb{R}^L$ and $\forall \mathbf{y} \in \mathbb{R}^N$. The cardinality of a countable set $\mathcal{T}$ is given by $\text{card}(\mathcal{T})$. The ℓ_p norm $\big(p \in]0, +\infty)\big)$ defined as

$$\|\mathbf{x}\|_p = \left(\sum_{n=1}^{N} (\mathbf{x}_n)^p \right)^{1/p}, \quad \forall \mathbf{x} \in \mathbb{R}^N$$

The ℓ_0 pseudo-norm [6], counting the non-zero entries of its argument, is defined as

$$(\forall \mathbf{x} \in \mathbb{R}^N) \quad \|\mathbf{x}\|_0 = \sum_{n=1}^{N} (\mathbf{x}_n)^0 ,$$

with the convention $0^0 = 0$. By abuse of notation, the ℓ_p norms and the ℓ_0 pseudo-norm will be used for tensors by reshaping them into vectors. Finally, the projection of a vector $\overline{\mathbf{x}} \in \mathbb{C}^N$ onto a non-empty closed subset $\mathcal{S}$ of $\mathbb{C}^N$ is given by

$$\mathcal{P}_{\mathcal{S}}(\overline{\mathbf{x}}) = \underset{\mathbf{x} \in \mathcal{S}}{\operatorname{argmin}} \|\mathbf{x} - \overline{\mathbf{x}}\|_2^2,$$

as denoted in [4].

1.2 MRI Physics

One of the main advantages of MRI is the ability to image soft tissue. This is possible due to the interaction of the applied magnetic fields with the proton spins of the soft tissues. In this section, the MRI magnetisation images will be presented as a solution to an inverse problem. A full description of the MRI experiment is not the main focus of this manuscript, for a more detailed description can be seen in [7–9].

1.2.1 Rotating Reference Frame

In order to mathematically describe the MRI experiment, two reference frames are considered. An absolute reference frame and a rotating reference frame that is used to simplify the description of the net magnetisation of the volume. By convention, the z axis (longitudinal axis) is defined with the applied static magnetic field $\mathbf{b}_0$. The x and y axis are defined as the orthogonal vectors in the plane normal to $\mathbf{b}_0$. Commonly, the y axis direction is defined from the floor to the ceiling and the x axis is chosen to follow the right-hand rule. The rotating reference frame is defined by rotating the absolute reference frame about the z axis with a non-zero angular frequency ω_0. Since

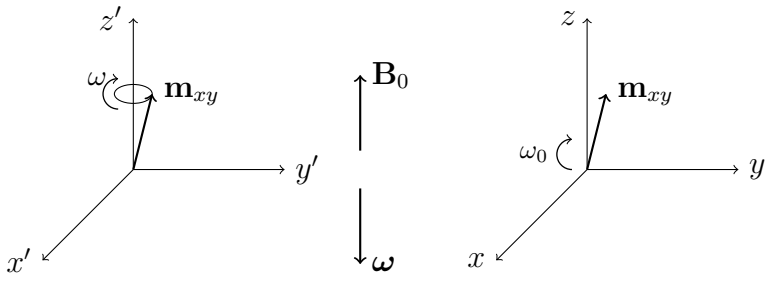


FIGURE 1.1: Spin precession seen from (left) the absolute reference frame (right) the rotating reference frame.

the spins of the protons will precess about $\mathbf{b}_0$ with their Larmor frequency $\omega = \gamma|\mathbf{b}_0|$ (approximately $\gamma = 42.6$MHz per Tesla), by setting $\omega_0 = \omega$ the spins appear to be stopped.

Let x', y' and z' be the three Cartesian axes in the absolute reference frame, and x, y and z the Cartesian axes in the rotating reference frame. Since $z' = z$, at time t, a vector in the absolute reference frame $\mathbf{p}'(t) \in \mathbb{R}^3$ is related to the rotating reference frame by:

$$\mathbf{p}(t) = \begin{bmatrix} \cos\omega_0 t & -\sin\omega_0 t & 0 \\ \sin\omega_0 t & \cos\omega_0 t & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{p}'(t) = \mathbf{R}\mathbf{p}'(t), \tag{1.3}$$

similarly,

$$\mathbf{p}'(t) = \begin{bmatrix} \cos\omega_0 t & \sin\omega_0 t & 0 \\ -\sin\omega_0 t & \cos\omega_0 t & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{p}(t) = \mathbf{R}^{-1}\mathbf{p}(t) = \mathbf{R}^T\mathbf{p}(t). \tag{1.4}$$

The relationship between the rates of change of a vector can be derived by taking the derivative with respect to t of (1.4) and translating the vector to the rotating reference frame:

$$\left(\frac{\partial \mathbf{p}(t)}{\partial t}\right)_{abs} = \left(\frac{\partial \mathbf{p}(t)}{\partial t}\right)_{rot} - \omega_0 \hat{\mathbf{z}} \times \mathbf{p}(t), \tag{1.5}$$

where the subscripts abs and rot represent the absolute reference frame and the rotational reference frame, and $\hat{\mathbf{z}} = [0\ 0\ 1]^T$. The net magnetisation of $\mathbf{m} = [m^x\, m^y\, m^z]^T$

of a group of spins precessing with the same phase in $\mathbf{b}_0$ can be described as:

$$\left(\frac{\partial \mathbf{m}}{\partial t}\right)_{abs} = \omega \mathbf{m} \times \hat{\mathbf{z}}. \tag{1.6}$$

Using equation (1.5), the rate of change of $\mathbf{m}$ in the rotating reference frame is:

$$\left(\frac{\partial \mathbf{m}}{\partial t}\right)_{rot} = \left(\frac{\partial \mathbf{m}}{\partial t}\right)_{abs} + \omega_0 \hat{\mathbf{z}} \times \mathbf{m} = (\omega - \omega_0) \times \hat{\mathbf{z}}. \tag{1.7}$$

If $\omega_0 = \omega$, $\mathbf{m}$ becomes a stationary vector in the rotating reference frame. The signal that generates the MRI signal is the transverse magnetisation. In order to produce such magnetisation, a radiofrequency (RF) magnetic field is introduced. Let $\mathbf{b}_1(t)$ be a RF magnetic field initially applied along the x axis and has an angular frequency of ω_{rf}. The $\mathbf{b}_1$ field in the absolute reference frame can be expressed as:

$$\mathbf{b}_1(t) = \begin{bmatrix} B_1(t)\cos\omega_{rf}t \\ -B_1(t)\sin\omega_{rf}t \\ 0 \end{bmatrix}. \tag{1.8}$$

In the rotating reference frame, if $\omega_{rf} = \omega$, the following expression is obtained:

$$\Big(\mathbf{b}_1(t)\Big)_{rot} = \mathbf{R}\mathbf{b}_1(t) = \begin{bmatrix} B_1(t) \\ 0 \\ 0 \end{bmatrix}, \tag{1.9}$$

The rotating frame has demodulated the RF oscillation and transforms the RF field in to the envelope $B_1(t)$.

1.2.2 Bloch equations

When only the $\mathbf{b}_0$ is applied, the spins align with this field and the net magnetisation at equilibrium, M_0 is proportional to the proton density ρ, within the voxel. Note that the equilibrium is not achieved instantaneously when $\mathbf{b}_0$ is applied, but it depends on the longitudinal relaxation time T_1 ($m^z(t) = M_0(1 - \exp(-t/T_1))$). If a transverse

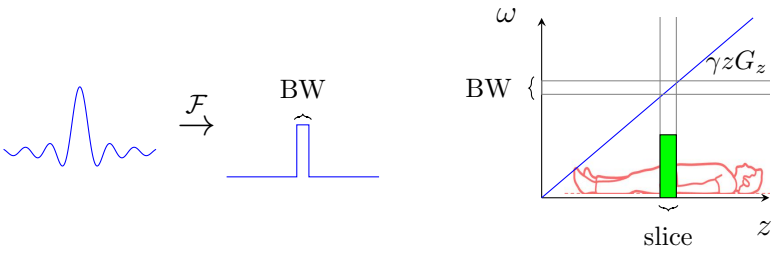


FIGURE 1.2: Example of the slice selection (Left) RF pulse in the time domain, (Centre) Fourier transform of the RF pulse, (Right) Corresponding image slice.

magnetisation exists $\mathbf{m}_{xy} = [m^x, m^y]$, it will decay as $|\mathbf{m}_{xy}(t)| = M_0 \exp(-t/T2)$. After equilibrium, the precession of the spins is manipulated applying $\mathbf{b}_1$ at the Larmor frequency. The overall magnetic field dynamics of the net magnetisation can be summarised by a set of differential equations called Bloch equations [7]:

$$\left(\frac{\partial \mathbf{m}(t)}{\partial t}\right)_{rot} = \mathbf{m}(t) \times \gamma \mathbf{b}(t) - \begin{bmatrix} m^x(t)/T_2 \\ m^y(t)/T_2 \\ (m^z - M_0)/T_1 \end{bmatrix}, \tag{1.10}$$

where $\mathbf{b} = \mathbf{b}_0 + \mathbf{b}_1$. The net magnetisation at a given time from an initial RF pulse can be obtained by integrating (1.10).

1.2.3 Spatial encoding and image formation

An image can be produced by spatially encoding the magnetisation signal while restricting the frequency band of the RF pulses. In order to select a slice of the volume, a gradient G_z is chosen to vary linearly in the z direction (i.e. $\mathbf{b}_0 = (G_z(z - z_0) + B_0)\hat{\mathbf{z}}$) as seen in Fig. 1.2. By applying an RF excitation pulse that only contain significant energy over a limited bandwidth (BW), corresponding to the Larmor frequencies in the selected slice. For small flip angles (generally, $\leq \pi/2$), following an RF pulse, $B_1(t)$, in the presence of a magnetic field gradient of amplitude G_z, the transverse magnetisation is given by:

$$|\mathbf{m}_{xy}| = \gamma M_0 \mathcal{F}\{B_1(t)\}_{f=\gamma G_z z/2\pi}, \tag{1.11}$$

where $\mathcal{F}$ is the one-dimensional (1D) Fourier transform at frequency f. If $\mathcal{F}\{B_1(t)\}$ has a rectangular distribution about ω_{rf}, only a rectangular distribution of spins around z_0 is tipped away from the z axis over a spatial region with a slice width of $2\pi\text{BW}/\gamma G_z$. If we assume that all the magnetisation initially lies along the z-axis, a rectangular profile can be achieved by setting $B_1(t) = \text{sinc}(\text{BW}t)$ (see Fig. 1.2). Note that $B_1(t)$ in practice has a finite duration, in consequence, the sinc function is not the optimal choice. Commonly $B_1(t)$ is tailored using Shinnar-Le Roux (SLR) algorithm [10]), this analysis is also used to explain when and how an excitation pulse can be used for refocusing [9]. Note that during this process, only the spins in the thin slice are tipped into the transverse plane (with a flip angle depending on $B_1(t)$).

The image formation is achieved using two spatially varying gradients G_x and G_y. In particular, let $\mathbf{b} = (B_0 + G_x x + G_y y)\hat{\mathbf{z}}$ be the magnetic field after the RF pulse excitation, assume that $|\mathbf{m}_{xy}|$ is relatively constant during the data acquisition and let the time at the centre of the acquisition be t_{acq}. The transverse magnetisation at each voxel at position (x, y) in the selected slice is given by:

$$m_{xy}(x, y, t) = m_{xy}(x, y, t_{acq}) \exp(-j\omega_{rf} t) \times \exp\left(-j\gamma \int_0^t (G_x(t')x + G_y(t')y)dt'\right), \tag{1.12}$$

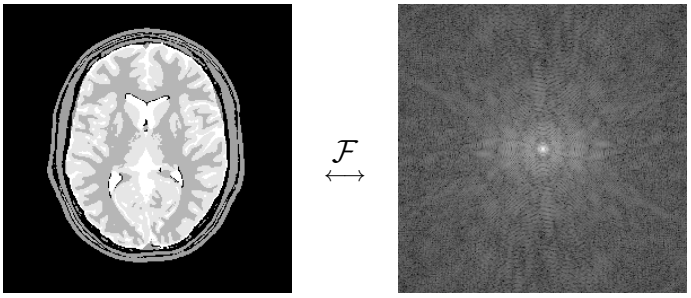
where $m_{xy} = m_x + jm_y$ is the complex representation of the vector $\mathbf{m}_{xy}$. By letting $k_x(t) = \gamma \int_0^t G_x(t')dt'$ and $k_y(t) = \gamma \int_0^t G_y(t')dt'$, the signal measured in a coil is modelled as:

$$v(t) = \exp(-j\omega_{rf} t) \int_x \int_y S(x, y) m_{xy}(x, y, t_{acq}) \times \exp\big((-j)(k_x(t)x + k_y(t)y)\big)dxdy, \tag{1.13}$$

where $S(x, y) \in \mathbb{C}$ is the coil sensitivity response at position (x,y). By demodulating the signal, the two-dimensional (2D) Fourier transform of $S(x, y)m_{xy}(x, y, t_{acq})$ at the frequencies $(k_x(t), k_y(t))$ is obtained:

$$s(t) = \exp(j\omega_{rf} t)v(t) = \mathcal{F}_2\big\{S(x, y)m_{xy}(x, y, t_{acq})\big\}\big(k_x(t), k_y(t)\big) \tag{1.14}$$

The (k_x, k_y) frequency domain, in the MRI community is known as k-space. Note



(a) (b)

FIGURE 1.3: Brain slice example (a) Spatial distribution of $|\mathbf{m}_{xy}|$ (b) Magnitude of the acquired signal in logarithmic scale.

that $s(t)$ is a complex function of t, in practice two separate signals are measured and combined to form the complex values. At a time instance, only single k-space point is acquired. An example of the transverse magnetisation magnitude of the imaged volume and the magnitude of the corresponding k-space can be seen in Fig. 1.3.

1.2.4 MRI as an inverse problem

Let $\mathbf{x} \in \mathbb{C}^N$ be the vector corresponding to the discretised spatial distribution of the transverse magnetisation with r_x voxels in x and r_y voxels in y ($N = r_x \times r_y$) and $\mathbf{y} \in \mathbb{C}^Q$ be the measurements given by (1.14), the acquisition model is simplified to:

$$\mathbf{y} = f(\mathbf{x}) + \boldsymbol{\eta}, \tag{1.15}$$

where f is the 2D non-uniform discrete Fourier transform (commonly this is implemented through its accelerated version the non-uniform fast Fourier Transform (NUFFT) in [11]), and $\boldsymbol{\eta} \in \mathbb{C}^Q$ is the noise in the acquisition governed by the Rician distribution. For acquisitions with the input signal-to-noise ratio (iSNR)>2dB, it is well approximated by a Gaussian distribution.

The non-uniform fast Fourier Transform for a trajectory $(k_x(t), k_y(t))$ with Q, can be expressed in a matrix form as:

$$f(\mathbf{x}) = \mathbf{G}\mathbf{x} \tag{1.16}$$

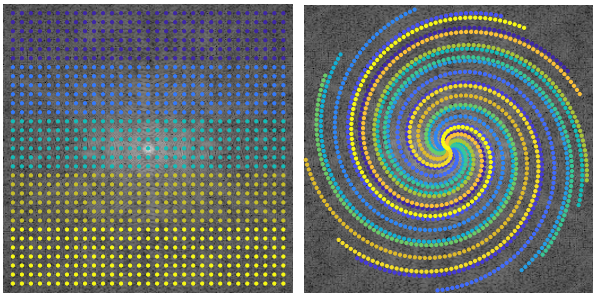


FIGURE 1.4: Examples of sampling patterns (a) Grid sampling (b) Spiral sampling.

where $\mathbf{G} \in \mathbb{C}^{Q \times N}$ is the non-uniform fast Fourier transform. The expression in (1.15), correspond to the forward model of the acquisition. Recovering the vector $\mathbf{x}$ from the $\mathbf{y}$ samples is called the inverse problem.

1.2.5 MRI challenges

The most popular reconstruction process to perform the 2D inverse Discrete Fourier Transform (2D-iDFT) to the full gridded k-space. In practice, the gridded k-space cannot be completely sampled after a single excitation. This is due to physical limitations, such as finite relaxation time of the dipoles and signal-to-noise ratio (SNR) limitations and technical limitations (variation limits of G_x and G_y). In consequence, the k-space is sampled in a sequence of L excitation-acquisition cycles with repetition time TR. In Fig. 1.4, examples of k-space sampling can be seen. The dots represent the acquired samples at a given cycle denoted by colour.

To overcome the time constraints of the acquisitions and based on the recent Compressed Sensing (CS) theory in [6, 12, 13], novel model-based methods have been proposed. These methods require less k-space samples to have accurate reconstructions, thus, the acquisition time can be significantly reduced. The computational realisation of these methods relies on concepts from optimisation theory.

Chapter 2

Optimisation tools

Mathematical optimisation theory can be used to obtain a solution to an inverse problem. We consider mathematical optimisation problems of the form:

$$\underset{\mathbf{x}\in\mathbb{R}^N}{\text{minimise}} \quad \sum_{i=1}^{m} f_i(\mathbf{x}) \tag{2.1}$$

where $f_i : \mathbb{R}^N \to \mathbb{R}$ [14]. Note that all the tools described in this chapter are defined for real numbers. In order to use the methods described in this chapter for complex numbers, we will consider:

$$\left(\forall \mathbf{x} \in \mathbb{C}^N\right) \to \begin{bmatrix} \text{real}\,(\mathbf{x}) \\ \text{imag}\,(\mathbf{x}) \end{bmatrix}. \tag{2.2}$$

The functions in (2.1) are used to describe what is a suitable solution for a given problem. A common criterion used to solve inverse problems is the least-squares criterion:

$$\underset{\mathbf{x}}{\text{minimise}} \quad \|h(\mathbf{x}) - \mathbf{y}\|_2^2, \tag{2.3}$$

where h is a function that describes the acquisition. The function in (2.3) is often called data fidelity term, and quantifies how close is the solution to the measurements. While in many applications solving (2.3) provides a good reconstruction, in particular, scenarios such as ill-posed problems (only partial information is available) the least-squares criterion amplifies the measurement noise. In these scenarios, including *prior* information in the optimisation problem has shown to improve the reconstructions.

Prior information is introduced through additional terms in the minimisation function. A common *prior* is to only allow solutions from a given set (i.e. only positive solutions). This kind of constraints can be modelled through indicator functions. Let

$\mathcal{T}$ be a non-empty set, the indicator function of $\mathcal{T}$ is defined as

$$\iota_{\mathcal{T}}(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \in \mathcal{T} \\ \infty & \text{otherwise.} \end{cases} \tag{2.4}$$

The projection $\mathcal{P}_{\mathcal{T}}(\mathbf{x})$ of $\mathbf{x} \in \mathbb{R}^N$ onto a non-empty set $\mathcal{T} \subset \mathbb{R}^N$ is defined as:

$$\mathcal{P}_{\mathcal{T}}(\mathbf{x}) = \underset{\overline{\mathbf{x}} \in \mathbb{R}^N}{\operatorname{argmin}} \quad \iota_{\mathcal{T}}(\overline{\mathbf{x}}) + \frac{1}{2}\|\mathbf{x} - \overline{\mathbf{x}}\|_2^2. \tag{2.5}$$

Note that (2.5), the projection $\mathcal{P}_{\mathcal{T}}(\mathbf{x})$ describes the closest element in $\mathcal{T}$ to $\mathbf{x}$ according to the ℓ_2-norm.

In this chapter, a brief review of the most popular optimisation algorithms to solve inverse problems is presented. In particular, algorithms that target two classes of optimisation problems are considered: the convex and the non-convex problems.

2.1 Convex Optimisation

In this section, a brief overview of the convex optimisation theory is presented. Convex optimisation problems have the same form as (2.1) and all the functions f_i are convex, which means the functions satisfy the inequality:

$$f_i(a\mathbf{x}_1 + (1-a)\mathbf{x}_2) \leq a f_i(\mathbf{x}_1) + (1-a) f_i(\mathbf{x}_2), \tag{2.6}$$

$\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^N$ and $0 \leq a \leq 1$. In a geometrical sense, this inequality describe that the line segment between $(\mathbf{x}_1, f_i(\mathbf{x}_1))$ and $(\mathbf{x}_2, f_i(\mathbf{x}_2))$ lies above the graph of f_i. A function is strictly convex if the strict inequality (2.6) holds whenever $\mathbf{x}_1 \neq \mathbf{x}_2$ [15]. Note that the minimiser of a convex problem is not always unique while for a strictly convex it is.

As an extension of the notion of the projection, the indicator function $\iota_{\mathcal{T}}$ can be replaced by any lower semicontinuous convex function f_c, the proximity operator of f_c is defined as

$$\text{prox}_{f_c}(\mathbf{x}) = \underset{\overline{\mathbf{x}} \in \mathbb{R}^N}{\text{argmin}} \quad f_c(\overline{\mathbf{x}}) + \frac{1}{2}\|\mathbf{x} - \overline{\mathbf{x}}\|_2^2. \tag{2.7}$$

The proximity operators are used in the iterative minimisation algorithms, due to their properties. Note that a large set of useful projections and proximity operators of convex functions have a closed-form solution (can be computed analytically). In the following subsections, the most popular algorithms to solve convex optimisation problems are described.

2.1.1 Gradient Descent

In the context of convex optimisation, gradient descent algorithm solves (2.1) with $m = 1$, where $f(\mathbf{x})$ is a convex and differentiable function with β-Lipschitz continuous gradient ∇f, i. e.:

$$\forall(\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^N \times \mathbb{R}^N) \ \|\nabla f(\mathbf{x}_1) - \nabla f(\mathbf{x}_2)\| \leq \beta\|\mathbf{x}_1 - \mathbf{x}_2\|, \tag{2.8}$$

where $\beta \in]0, \infty[$. By iterating:

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} - \mu\nabla f(\mathbf{x}^{(i)}), \tag{2.9}$$

where $\mu < 2/\beta$, the iterates will converge to a solution of the minimisation problem [14, 16, 17].

2.1.2 Forward-Backward Splitting

The forward-backward algorithm solves the (2.1) with m=2, where $f_1(\mathbf{x})$ is a convex and differentiable function with β-Lipschitz continuous gradient ∇f_1 and $f_2(\mathbf{x})$ is a

convex and non-smooth function. The details can be seen in Algorithm 2.

Algorithm 2 Forward-backward algorithm

1: **Iterations:**
2: **for** $i = 0, 1, \ldots$ **do**
3: **Forward Step:**
4: $\overline{\mathbf{x}}^{(i)} = \mathbf{x}^{(i)} - \mu^{(i)} \nabla f_1 \left(\mathbf{x}^{(i)} \right)$
5: **Backward Step:**
6: $\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \lambda^{(i)} \left(\text{prox}_{\mu^{(i)} f_2} \left(\overline{\mathbf{x}}^{(i)} \right) - \mathbf{x}^{(i)} \right)$

The sequence generated by Algorithm 2 converges to a solution of the problem if $0 < \mu^{(i)} < 2/\beta$ and $0 < \lambda^{(i)} < 1$ [14, 18]. Note that the choice of $\mu^{(i)}$ and $\lambda^{(i)}$ affects the convergence rate.

If f_2 is the indicator function of a convex set $\mathcal{C}$ and $\lambda^{(i)}$ is fixed to be 1, the forward-backward algorithm reduces to the projected gradient algorithm:

$$\mathbf{x}^{(i+1)} = \mathcal{P}_{\mathcal{C}} \left(\mathbf{x}^{(i)} - \mu^{(i)} \nabla f_1 \left(\mathbf{x}^{(i)} \right) \right). \tag{2.10}$$

2.2 Non-convex Optimisation

As discussed in the previous section, in a convex optimisation problem all the involved functions are convex. A non-convex minimisation problem has the same form as (2.1) with at least one function that does not satisfy (2.6). In this manuscript, only a particular case of a non-convex problem is considered:

$$\underset{\mathbf{x} \in \mathcal{N}}{\text{minimise}} \quad f(\mathbf{x}) \tag{2.11}$$

where $\mathcal{N} \subset \mathbb{R}^N$ is a non-convex set and f is a differentiable function. In contrast to convex optimisation algorithms, non-convex optimisation algorithms can only guarantee global optimality under specific assumptions and depending on the functions involved.

2.2.1 Forward-Backward Algorithm

In the particular case of (2.11), the forward-backward algorithm is equivalent to the projected gradient algorithm in the non-convex setting:

$$\mathbf{x}^{(i+1)} = \mathcal{P}_{\mathcal{N}}\left(\mathbf{x}^{(i)} - \mu^{(i)}\nabla f\left(\mathbf{x}^{(i)}\right)\right), \tag{2.12}$$

where,

$$\mathcal{P}_{\mathcal{N}}\left(\mathbf{x}\right) = \underset{\overline{\mathbf{x}}\in\mathcal{N}}{\operatorname{argmin}} \|\mathbf{x} - \overline{\mathbf{x}}\|_2^2. \tag{2.13}$$

Note that the projection in (2.13) can be an NP-hard problem. In particular cases, this kind of projection can be computed approximately in feasible time.

The sequence generated by the algorithm in (2.12) is guaranteed to converge to a local minimum [19]. Under specific conditions of $\mathcal{N}$, f and $\mu^{(i)}$, the algorithm in (2.12) is able to find the global solution of (2.11).

2.2.2 Block Gauss-Seidel method

Consider the following minimisation problem:

$$\underset{\mathbf{x}_1\in\mathcal{C}_1,\mathbf{x}_2\in\mathcal{C}_2}{\operatorname{argmin}} f(\mathbf{x}_1, \mathbf{x}_2), \tag{2.14}$$

where f is a strict quasi-convex with respect to the blocks $\mathbf{x}_1$ and $\mathbf{x}_2$, and $\mathcal{C}_1$ and $\mathcal{C}_2$ are convex sets. A function is strictly quasi-convex if it satisfies with respect to the block $\mathbf{x}_1$:

$$f\left(a\mathbf{x}_1 + (1-a)\hat{\mathbf{x}}_1, \mathbf{x}_2\right) < \max\left\{f(\mathbf{x}_1, \mathbf{x}_2), f(\hat{\mathbf{x}}_1, \mathbf{x}_2)\right\} \tag{2.15}$$

$\forall \mathbf{x}_1, \hat{\mathbf{x}}_1, \mathbf{x}_2 \in \mathbb{R}^N$ and $0 \leq a \leq 1$. By alternating:

$$\mathbf{x}_1^{(i+1)} = \underset{\mathbf{x}_1\in\mathcal{C}_1}{\operatorname{argmin}} f(\mathbf{x}_1, \mathbf{x}_2^{(i)}), \tag{2.16}$$

and

$$\mathbf{x}_2^{(i+1)} = \underset{\mathbf{x}_2\in\mathcal{C}_2}{\operatorname{argmin}} f(\mathbf{x}_1^{(i+1)}, \mathbf{x}_2), \tag{2.17}$$

the algorithm is guaranteed to converge into a critical point of problem (2.14) [20].

2.2.3 Block Coordinate Forward-Backward Algorithm

Consider the following minimisation problem:

$$\underset{\mathbf{x}_1 \in \mathcal{C}_1, \mathbf{x}_2 \in \mathcal{C}_2}{\operatorname{argmin}} \; f(\mathbf{x}_1, \mathbf{x}_2) + g_1(\mathbf{x}_1) + g_2(\mathbf{x}_2), \tag{2.18}$$

where f is a strict quasi-convex differentiable (see (2.15)) function with respect to the blocks $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathcal{C}_1$ and $\mathcal{C}_2$ are convex sets and g_1 and g_2 are convex functions.

Algorithm 3 Block Coordinate Forward-Backward Algorithm

Input: $\mathbf{x}_1^{(0)} \in \mathcal{C}_1$, $\mathbf{x}_2^{(0)} \in \mathcal{C}_2$, $(I, J) \in \mathbb{N}_*^2$, $(\mu, \nu) \in (\mathbb{R}_*^+)^2$
for $k = 0, 1, \ldots$ **do**
  $\widetilde{\mathbf{x}}_1^{(0)} = \mathbf{x}_1^{(k)}$ and $\widetilde{\mathbf{x}}_2^{(0)} = \mathbf{x}_2^{(k)}$
  for $i = 0, 1, \ldots, I-1$ **do**
    $\widetilde{\mathbf{x}}_1^{(i+1)} = \mathcal{P}_{\mathcal{C}_1}\left(\operatorname{prox}_{\mu g_1(\cdot)}\left(\widetilde{\mathbf{x}}_1^{(i)} - \mu \nabla_{\mathbf{x}_2} f\left(\widetilde{\mathbf{x}}_1^{(i)}, \mathbf{x}_2^{(k)}\right)\right)\right)$
  $\mathbf{x}_1^{(k+1)} = \widetilde{\mathbf{x}}_1^{(I)}$
  for $j = 0, 1, \ldots, J-1$ **do**
    $\widetilde{\mathbf{x}}_2^{(j+1)} = \mathcal{P}_{\mathcal{C}_2}\left(\operatorname{prox}_{\nu g_2(\cdot)}\left(\widetilde{\mathbf{x}}_2^{(j)} - \nu \nabla_{\mathbf{x}_1} f\left(\mathbf{x}_1^{(k+1)}, \widetilde{\mathbf{x}}_2^{(j)}\right)\right)\right)$
  $\mathbf{x}_2^{(k+1)} = \widetilde{\mathbf{x}}_2^{(J)}$

Under technical conditions described in [3], the sequence $(\mathbf{x}_1^{(k)}, \mathbf{x}_2^{(k)})_{k \in \mathbb{N}}$ generated by Algorithm 9 is guaranteed to converge to a critical point of the objective function (2.18). Note that $\nabla_{\mathbf{x}_1} f$ (resp. $\nabla_{\mathbf{x}_2} f$) denotes the partial gradient of f with respect to $\mathbf{x}_1$ (resp. $\mathbf{x}_2$)

Chapter 3

Magnetic Resonance Fingerprinting

As stated in Chapter 1, MRI aims to recover the spatial distribution of the transverse magnetisation of a given volume. However, these reconstructions are qualitative or weighted measurements of a limited set of properties (i.e., relaxation times, proton density). The same kind of material can have different intensities in different acquisitions depending on several factors such as the scanner configuration, inhomogeneities of the magnetic fields and timing errors.

MRI acquisitions that aim to quantify the set of properties is called quantitative MRI (qMRI). Traditional qMRI methods acquire several MRI images in such a way that the parameters can be extracted by a correlation of these images. While today these methods are still the common choice in qMRI, prohibitively long acquisition times and in consequence inaccuracies in the reconstructions do not allow this kind of studies to become a standard for diagnosis.

In recent years, a growing number of papers attempt to address the acquisition problem using CS theory such as [21–23]. These techniques accelerate the acquisition by a partial sampling of the k-space and the fitting of a decaying exponential to retrieve the properties of interest. Nonetheless, the main focus of these methods is the development of advanced reconstructions algorithms and do not address the fundamental issue of how to design the experiment for an efficient acquisition.

A novel acquisition was recently presented in [2] that allows simultaneous quantification of multiple tissue properties through a single process. This method is called Magnetic Resonance Fingerprinting (MRF). The novelty of this acquisition is that there is no need for the material to return to equilibrium between the excitation pulses and in a combination of k-space subsampling schemes the acquisition process

is significantly reduced. The pulse sequences are used to encode multiple parameters. A set of predicted signals (through the Bloch equations or the extended phase graphs (EPG) model) is used to retrieve the parameters of interest from the measured data using a matching filter.

This chapter is structured as follows: the first part is the MRF model as an inverse problem, the second part is a review of the state-of-the-art algorithms and in the last part a novel algorithms for accelerated iterative reconstruction that allow super-resolution reconstructions relying on spatial regularisation are proposed.

3.1 MRF Model

The MRF acquisition process is based on a sequence of excitation pulses, depending on the sequence the measured signal is only sensitive to a particular set of parameters corresponding to the chemical composition of the volume of interest. The design of these sequences is out of the scope of this manuscript. While the optimal choice of the acquisition parameter is still an open question a theoretical analysis is provided in [24] for the sequences based in the inversion recovery steady-state free precession (IR-SSFP) pulse sequence.

Let $\mathbf{Y} \in \mathbb{C}^{Q\times L\times C}$ be the measurement matrix, where L is the excitation sequence length, C is the number of coils and Q is the number of measurements at each excitation and each coil. Let $\mathbf{M} \in \mathbb{C}^{N\times L}$ be the transverse magnetisation response of the imaged volume of interest with N voxels. For every $(l, c) \in \{1, \dots, L\} \times \{1, \dots, C\}$, the corresponding observation $\mathbf{Y}_{:,l,c} \in \mathbb{C}^{Q}$ is given by

$$\mathbf{Y}_{:,l,c} = \mathbf{G}_l \mathrm{Diag}\left(\mathbf{S}_{:,c}\right) \mathbf{M}_{:,l} + \boldsymbol{\eta}_{:,l,c}, \tag{3.1}$$

where $\mathbf{G}_l$ is the non-uniform Fourier transform with the k-space trajectory $(k_x^{(l)}, k_y^{(l)})$, $\mathbf{S} \in \mathbb{C}^{N\times C}$ is the concatenation of C spatial sensitivity coil maps, and $\boldsymbol{\eta} \in \mathbb{C}^{Q\times L\times C}$ is a realisation of a random i.i.d. Gaussian noise (Note that this is an approximation of

the Rician noise). An example of the measurements $\mathbf{Y}_{:,l,c}$ is shown in Fig. 3.1. Note that the figure is an illustration of the vectorised operations in (3.1).

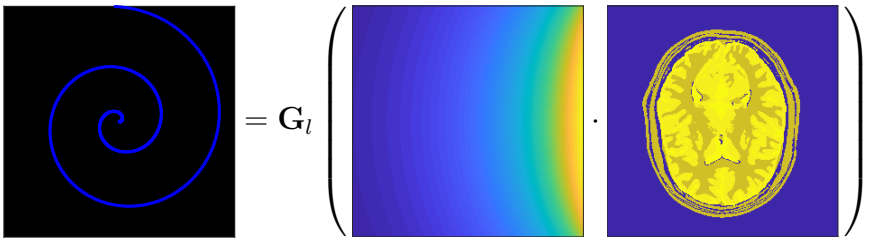


FIGURE 3.1: Example of a k-space sampling for an acquisition l and coil c.

Let $h : \mathbb{C}^{N \times L} \to \mathbb{C}^{Q \times L \times C}$ be the linear mapping defining the complete acquisition process such that:

$$\mathbf{Y} = h\left(\mathbf{M}\right) + \boldsymbol{\eta}. \tag{3.2}$$

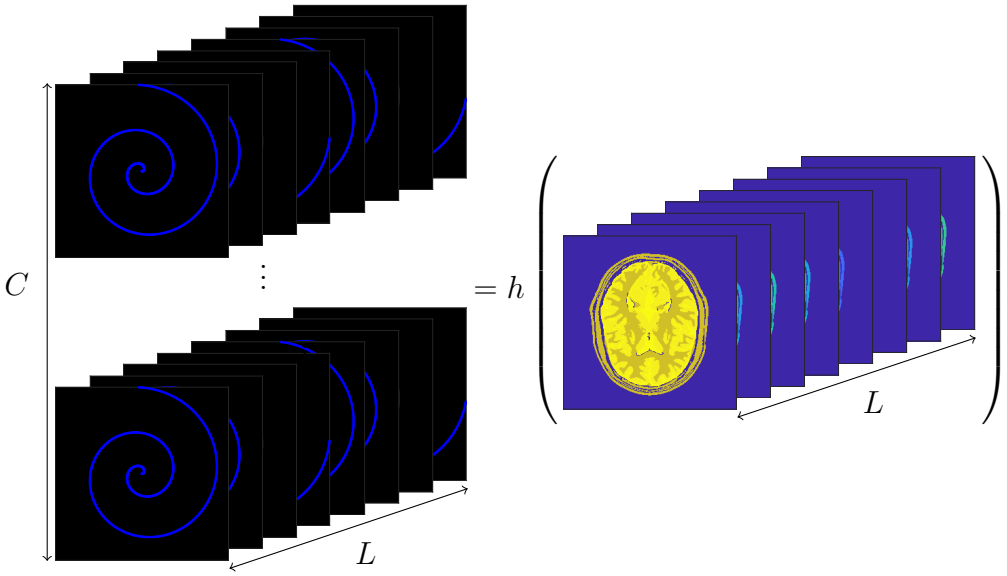


FIGURE 3.2: Example of the measurement operator.

An example of the operator h can be seen in Fig. 3.2. For each voxel $n \in \{1, \ldots, N\}$, the magnetisation response $\mathbf{M}_{n,:}$ is modelled through the smooth non-linear mapping $B : \mathcal{M} \to \mathbb{C}^{1 \times L}$ scaled by the unknown proton density $\rho_n \in \mathbb{R}_+$:

$$\mathbf{M}_{n,:} = \rho_n B(\hat{\boldsymbol{\theta}}_{n,:}, \boldsymbol{\Gamma}) \tag{3.3}$$

where $\mathbf{\Gamma} \in \mathbb{R}^{A\times 1}$ represents the concatenation of A known acquisition parameters (*e.g.*, flip angles α, repetition times TR) chosen such that $\mathbf{M}_{n,:}$ is only sensitive to the P parameters $\hat{\boldsymbol{\theta}}_{n,:} \in \mathcal{M}$ under investigation, where $\mathcal{M} \subset \mathbb{R}^{1\times P}$ denotes the subset of feasible parameters. In the reminder of this manuscript, $P = 2$ is fixed and corresponds to the relaxation times T_1 and T_2. Note that the B can be approximated by the Bloch equations or more realistic models such as the Extended Phase Graphs (EPG) formalism in [25].

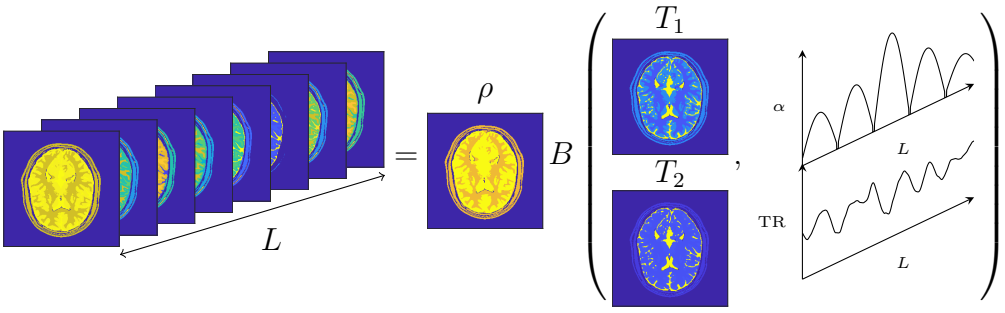


FIGURE 3.3: Example of the non-linear operator B.

3.2 MRF Reconstruction Techniques

In this section, the state-of-the-art methods for MRF are presented. In particular, the matching filter proposed in [2], the SVD temporal compression proposed in [26], the iterative reconstructions proposed in [27, 28] and regularised methods.

3.2.1 Matching Filter

The key idea behind the MRF acquisition is to produce unique temporal patterns (*fingerprints*) $\mathbf{M}_{n,:}$ in such a way that even with the high aliasing produced by the subsampling and the uncorrelated noise introduced in the acquisition, the parameter information is preserved.

The reconstruction method proposed in [2] can be summarised in two steps: reconstruction of the aliased noisy magnetisation images and a voxel-wise comparison of

the magnetisation evolution to the predicted signals through a matched filter. More precisely, the magnetisation images are reconstructed as follows:

$$\overline{\mathbf{M}} = h^{\dagger}\left(\mathbf{Y}\right), \tag{3.4}$$

note that because of the k-space subsampling and the noise acquisition $\overline{\mathbf{M}}$ is imperfect. To avoid dealing with the function B in (3.3), the authors proposed to create an over-complete dictionary $\mathbf{\Phi} \in \mathbb{C}^{D\times L}$ of fingerprints, as a discrete sampling of the low dimensional manifold B. Then, $\mathbf{\Phi}$ is constructed from D samples of $\mathcal{M}$, stored in a matrix $\boldsymbol{\theta} \in \mathbb{R}^{D\times P}$. An example of the dictionary can be seen in Fig. 3.4.

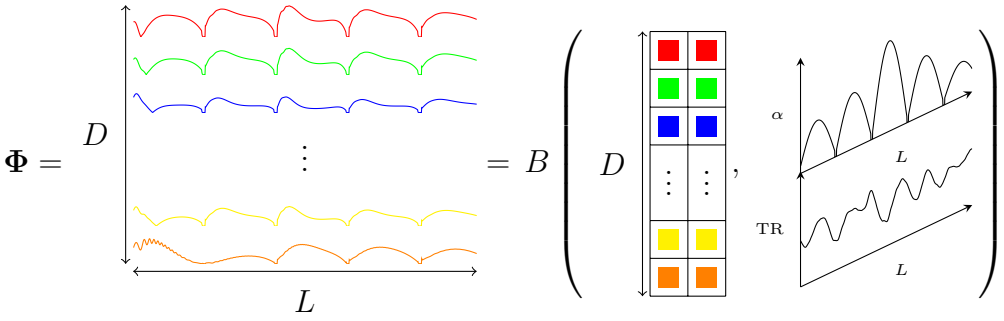


FIGURE 3.4: Example of the fingerprint dictionary.

Each entry of the dictionary $\mathbf{\Phi}$, is compared to the estimated fingerprint and the parameters of the best match by the cosine similarity criteria (inner product):

$$\overline{d}_n = \underset{d}{\operatorname{argmax}} \quad \frac{\operatorname{real}\left(\overline{\mathbf{M}}_{n,:}\mathbf{\Phi}^{\dagger}_{d,:}\right)}{\|\mathbf{\Phi}_{d,:}\|_2}, \tag{3.5}$$

the matrix $\boldsymbol{\theta}$ is used as a Look-up table (LUT) to retrieve the parameters of interest. The proton density is estimated using the vector projection of the fingerprint to the estimated dictionary element:

$$\overline{\rho}_n = \max\left(\frac{\operatorname{real}\left(\overline{\mathbf{M}}_{n,:}\mathbf{\Phi}^{\dagger}_{\overline{d}_n,:}\right)}{\|\mathbf{\Phi}_{\overline{d}_n,:}\|_2^2}, 0\right). \tag{3.6}$$

While the proton density is positive by definition, in practice it is allowed to be complex-valued to compensate for model errors such as timing and calibration errors.

The matching process for complex proton densities is:

$$\overline{d}_n = \underset{d}{\operatorname{argmax}} \quad \frac{\left|\overline{\mathbf{M}}_{n,:}\mathbf{\Phi}^{\dagger}_{d,:}\right|}{\|\mathbf{\Phi}_{d,:}\|_2}, \tag{3.7}$$

and,

$$\overline{\rho}_n = \frac{\left|\overline{\mathbf{M}}_{n,:}\mathbf{\Phi}^{\dagger}_{\overline{d}_n,:}\right|}{\|\mathbf{\Phi}_{\overline{d}_n,:}\|_2^2}. \tag{3.8}$$

For a given acquisition process, the reconstruction quality mainly depends on the quality of the magnetisation images and how fine the discretisation of the manifold is. Increasing the number of dictionary elements (finer discretisation) can introduce computational problems such as high memory requirements and high reconstruction times. Note that the reconstruction process has no explicit link to CS theory and do not provide any recovery guarantees.

3.2.2 SVD Temporal Compression

As mentioned before, the reconstruction process highly depends on the manifold discretisation. For accurate reconstructions, huge dictionaries ($D \sim 10^6$) are used, significantly increasing the reconstruction time. To address this problem, some methods accelerate the matching process by exploiting the correlation of the dictionary elements [29, 30]. To further accelerate the reconstructions, a compression method on the temporal dimension L was proposed in [26] using the singular-value decomposition of the dictionary:

$$\mathbf{\Phi} = \mathbf{W}\mathbf{E}\mathbf{V}^{\dagger}, \tag{3.9}$$

where $\mathbf{W} \in \mathbb{C}^{D\times D}$ and $\mathbf{V} \in \mathbb{C}^{L\times L}$ are unitary matrices, and $\mathbf{E}$ is the diagonal matrix containing the singular values. If over-complete dictionary $\mathbf{\Phi}$ contain all the elements that describe the transverse magnetisation of the volume, then the magnetisation can be expressed as:

$$\mathbf{M} = \mathbf{X}\mathbf{\Phi}, \tag{3.10}$$

where $\mathbf{X} \in \mathbb{R}_{+}^{N \times D}$ is the mixing matrix containing the proton densities in the positions of the corresponding dictionary elements (since only a dictionary element per voxel is expected, each row contains at most one value different than 0). The equation (3.1) can be rewritten as:

$$\mathbf{Y}_{:,l,c} = \Omega\left(\mathbf{G}\text{Diag}\left(\mathbf{S}_{:,c}\right)\mathbf{M}\right) + \boldsymbol{\eta}_{:,:,c} \tag{3.11}$$

where $\Omega : \mathbb{C}^{K \times L} \rightarrow \mathbb{C}^{Q \times L}$ is the selection operator that for each excitation instance $l \in \{1, \ldots, L\}$, takes the corresponding measurements and $\mathbf{G}$ is the non-uniform fast Fourier transform of the K points corresponding to the combined trajectories $(k_x^{(l)}, k_y^{(l)})$. By combining (3.9) and (3.10), into (3.11), the following expression can be obtained:

$$\mathbf{Y}_{:,:,c} = \Omega\left(\mathbf{G}\text{Diag}\left(\mathbf{S}_{:,c}\right)\mathbf{X}\mathbf{W}\mathbf{E}\mathbf{V}^{\dagger}\right) + \boldsymbol{\eta}_{:,:,c}. \tag{3.12}$$

Let $\widehat{\mathbf{E}} \in \mathbb{C}^{D \times V}$ be the truncated version of $\mathbf{E}$ and $\widehat{\mathbf{V}} \in \mathbb{C}^{L \times V}$ be the corresponding singular vectors of $\mathbf{V}$, the dictionary can be approximated by:

$$\mathbf{W}\mathbf{E}\mathbf{V}^{\dagger} \approx \mathbf{W}\widehat{\mathbf{E}}\widehat{\mathbf{V}}^{\dagger}, \tag{3.13}$$

thus, the measurements can be approximated by,

$$\mathbf{Y}_{:,:,c} \approx \Omega\left(\mathbf{G}\text{Diag}\left(\mathbf{S}_{:,c}\right)\mathbf{X}\mathbf{W}\widehat{\mathbf{E}}\widehat{\mathbf{V}}^{\dagger}\right) + \boldsymbol{\eta}_{:,:,c}. \tag{3.14}$$

Since,

$$\mathbf{M}\widehat{\mathbf{V}} = \mathbf{X}\mathbf{W}\mathbf{E}\mathbf{V}^{\dagger}\widehat{\mathbf{V}} = \mathbf{X}\mathbf{W}\widehat{\mathbf{E}}, \tag{3.15}$$

then,

$$\mathbf{Y}_{:,:,c} \approx \Omega\left(\left[\mathbf{G}\text{Diag}\left(\mathbf{S}_{:,c}\right)\mathbf{M}\widehat{\mathbf{V}}\right]\widehat{\mathbf{V}}^{\dagger}\right) + \boldsymbol{\eta}_{:,:,c}. \tag{3.16}$$

Note that by performing the operations in the brackets first the complexity is significantly reduced (i.e. only $V << L$ NUFFT operations are required). As described in [26], further acceleration can be achieved by working with the compressed version of the magnetisation sequence $\widehat{\mathbf{M}} = \mathbf{M}\widehat{\mathbf{V}}$. Let $h_V : \mathbb{C}^{N \times V} \rightarrow \mathbb{C}^{Q \times L \times C}$ then, the equation (3.16) can be expressed as:

$$\mathbf{Y} \approx h_V\left(\widehat{\mathbf{M}}\right) + \boldsymbol{\eta}. \tag{3.17}$$

Similarly to [2], the reconstruction process is performed in two steps, reconstructing a dirty version of the compressed magnetisation images using:

$$\overline{\mathbf{M}} = h_V^\dagger \left(\mathbf{Y} \right), \tag{3.18}$$

and then matching each compressed fingerprint to an element of the compress dictionary $\widehat{\mathbf{\Phi}} = \mathbf{\Phi}\widehat{\mathbf{V}}$ using equations (3.5) and (3.6), for positive proton densities or equations (3.7) and (3.8), for complex proton densities. Note that $h_V^\dagger$ is significantly less expensive than $h^\dagger$ (since only V NUFFT operations are needed) and the matching process is performed using shorter fingerprints. An important remark is that the quality of the approximation depends on how fast is the decay of the singular values. In [26], it was shown that the MRF signal can be compressed significantly allowing accelerated reconstructions without relevant accuracy loss. While this method addresses the computational issues of the reconstruction process, the conditions for a stable recovery are not provided.

3.2.3 Iterative algorithms

More recently, a full CS strategy was formulated in [27]. In this work, the authors proposed to solve:

$$\underset{\mathbf{M} \in \mathcal{B}^+}{\text{minimise}} \ \frac{1}{2} \| \mathbf{Y} - h\left(\mathbf{M} \right) \|_2^2, \tag{3.19}$$

where the non-convex set $\mathcal{B}^+$ is defined as:

$$\mathcal{B}^+ = \Big\{ \mathbf{M} \in \mathbb{C}^{N \times L} \,|\, (\forall n \in \{1, \ldots, N\}) \ \mathbf{M}_{n,:} = \rho \mathbf{m}, \\ \text{with } \rho \in \mathbb{R}_+ \text{ and } \mathbf{m} \in B\left(\mathcal{M}, \mathbf{\Gamma} \right) \Big\}. \tag{3.20}$$

In [27], an iterative projection algorithm called BLoch recovery via Iterative Projection (BLIP), based on (2.12), was proposed to solve (3.19). This method imposes that $\mathbf{M}$ belongs to $\mathcal{B}^+$ by computing at iteration i:

$$\mathbf{M}^{(i+1)} = \mathcal{P}_{\mathcal{B}^+} \left(\mathbf{M}^{(i)} - \mu^{(i)} h^\dagger \left(h\left(\mathbf{M}^{(i)} \right) - \mathbf{Y} \right) \right), \tag{3.21}$$

where $\mu^{(i)} > 0$ is a step size chosen using a backtracking method and $h^\dagger$ is the adjoint operator of h, and $\mathcal{P}_{\mathcal{B}^+}$ represents the projection onto the set $\mathcal{B}^+$. The authors also derived a technical condition on both L and the undersampling ratio N/Q, for the Restricted Isometry Property (RIP) to be satisfied by the operator h which guarantees robust recovery. In this work, the authors interpreted the dictionary $\mathbf{\Phi}$ of fingerprints, introduced in [2], as a discrete sampling of the low dimensional manifold B, and the matched filter as the projection $\mathcal{P}_{\mathcal{B}^+}$. While the results were obtained for an Echo-planar Imaging (EPI) subsampling scheme [31], BLIP has shown to behave well different k-space trajectories.

In [27], it was shown that the BLIP method outperforms the matching method from Section 3.2.1 for short acquisitions (smaller L). This is expected since the matching method proposed in [2] can be seen as the first iteration of the BLIP method with a step size of $\mu^{(1)} = 1$ and $\mathbf{M}^{(0)}$ is the null matrix.

An accelerated version of BLIP using the SVD compression discussed in the previous subsection was proposed in [28] called Accelerated Iterative Reconstruction for MRF (AIR-MRF). Similarly to BLIP method, AIR-MRF aims to solve:

$$\underset{\mathbf{M}\in\mathcal{B}_V^+}{\text{minimise}} \ \frac{1}{2}\left\|\mathbf{Y} - h_V\left(\widehat{\mathbf{M}}\right)\right\|_2^2, \tag{3.22}$$

where,

$$\mathcal{B}_V^+ = \left\{\widehat{\mathbf{M}} \in \mathbb{C}^{N\times V} \,|\, (\forall n \in \{1,\dots,N\})\ \widehat{\mathbf{M}}_{n,:} = \rho\mathbf{m}, \right. \\ \left. \text{with } \rho \in \mathbb{R}_+ \text{ and } \mathbf{m} \in B\left(\mathcal{M}, \mathbf{\Gamma}\right)\widehat{\mathbf{V}}\right\}. \tag{3.23}$$

using BLIP method in the compressed temporal space:

$$\widehat{\mathbf{M}}^{(i+1)} = \mathcal{P}_{\mathcal{B}_V^+}\left(\widehat{\mathbf{M}}^{(i)} - \mu^{(i)} h_V^\dagger\left(h_V\left(\widehat{\mathbf{M}}^{(i)}\right) - \mathbf{Y}\right)\right). \tag{3.24}$$

As mentioned in [28], depending on the k-space sampling pattern $(k_x^{(l)}, k_y^{(l)})$, the reconstructions of (3.21) and (3.24) could produce artefacts. In particular, the spiral pattern used in [2] and [26] introduce high frequency artefacts.

3.2.4 Spatial Regularisation

In order to reduce the effect of the sampling artefacts, the authors of [28] proposed to incorporate a spatial regularisation term into by incorporation a low-pass filter in the reconstruction process. The authors propose a filter constructed by two concentric circles, the outer circle is defined to match the highest sampled frequency and the inner circle to the 85% of the highest sampled frequency. The filter varies linearly from 0 to 1 from the outer circle to the inner circle. The algorithm can be expressed as:

$$\widehat{\mathbf{M}}^{(i+1)} = g\left(\mathcal{P}_{\mathcal{B}_V^+}\left(\widehat{\mathbf{M}}^{(i)} - \mu^{(i)} h_V^\dagger\left(h_V\left(\widehat{\mathbf{M}}^{(i)}\right) - \mathbf{Y}\right)\right)\right). \tag{3.25}$$

where g is a low pass filter applied to the spatial distribution of the proton density. While this kind of approach seems to reduce the artefacts of the reconstructions, the authors do not give any insights on the convergence of the algorithm or how it could be linked to a minimisation problem.

3.3 Proposed Methods

As discussed in the previous subsection the methods in [2, 26] can be seen as the first iteration of the methods in [27, 28]. While iterative algorithms allow reconstructions with significantly fewer measurements, the reconstruction time is increased. AIR-MRF accelerates the reconstruction by compressing the temporal dimension, while the matching procedure is accelerated due to the shorter fingerprints, the number of dictionary fingerprints remain the same. For high resolved volumes and/or high reconstruction quality the number of dictionary fingerprints D could significantly increase the reconstruction time. In this section, we propose methods to accelerate the projection step and allow super-resolution reconstructions by incorporating spatial regularisation.

3.3.1 Adaptive-BLIP

Commonly MRF aims to obtain quantitative values of a small set of tissues. In practice, only $T \ll D$ elements of the dictionary $\mathbf{\Phi}$ are necessary to characterise $\mathbf{M}$. While T is unknown, we have a reasonable estimate for it. A loose upper bound K is introduced, such that $T \leq K \leq D$, to limit the number of active dictionary elements. Taking advantage of the smoothness of B, the dictionary can be refined through iterations. These process can be seen as an approximate projection and will be discussed in detail in the next chapter.

In this manuscript, the dictionary is updated using the following procedure. A coarse dictionary is first defined using a fixed grid. After the matched filter, quantitative parameters $\boldsymbol{\theta}_c$ are clustered by the K-means algorithm based on the fingerprint distribution across all voxels. The number of clusters K can be defined as proportional to the number of expected tissues in the volume to be imaged. The dictionary is updated $\mathbf{\Phi}^{(i)}$ to refine the manifold elements of interest. For this process, n_s random samples are generated around the clusters $\boldsymbol{\theta}_c$ using a Gaussian distribution with a diagonal covariance matrix $\mathbf{\Sigma}^{(i)}$. The values of the covariance matrix $\mathbf{\Sigma}^{(i)}$ are reduced by a factor $0 < \beta < 1$ at each iteration. When the values of $\mathbf{\Sigma}^{(i)}$ are sufficiently small, the used elements in dictionary $\mathbf{\Phi}^{(i)}$ will not change anymore and after a fixed number of iterations, the sequences generated by Algorithm 4 will stabilise. Since the samples are randomly Gaussian distributed, the parameter values are not limited to a given resolution. Note that the projection using the dictionary $\mathbf{\Phi}^{(i)}$ is an approximation. This algorithm is dubbed Adaptive-BLIP (ABLIP) and it was presented at BASP Frontiers Workshop 2017 in Switzerland.

3.3.2 Spatial regularisation

As discussed in the previous section, the sampling can introduce artefacts in the reconstructions. We propose to regularise our solution by adding an additional term

Algorithm 4 Adaptive-BLIP

Input: $\mathbf{Y} \in \mathbb{C}^{Q\times L\times C}$, $\zeta < 1$, $\mathbf{M}^{(0)} \in \mathbb{C}^{N\times L}$
Iterations:
for $i = 0, 1, \ldots$ **do**
  $\mu = 2N/Q$, $\nu = 0$
  while $\mu > \nu$ **do**
    $\mu = \mu/2$
    Gradient Step:
    $\overline{\mathbf{M}}^{(i)} = \mathbf{M}^{(i)} - \mu h^{\dagger}\left(h\left(\mathbf{M}^{(i)}\right) - \mathbf{Y}\right)$
    Dictionary Estimation:
    $\mathbf{\Phi^{(i+1)}} = \text{DictionaryEstimation}\left(\overline{\mathbf{M}}^{(i)}, \mathbf{\Phi}^{(i)}\right)$
    Projection Step:
    $\mathbf{M}^{(i)} = \mathcal{P}_{\mathcal{B}^{+}\left(\mathbf{\Phi}^{(i+1)}\right)}\left(\overline{\mathbf{M}}^{(i)}\right)$
    Backtracking step:
    $\nu = \zeta \frac{\|\mathbf{M}^{(i+1)} - \mathbf{M}^{(i)}\|_2^2}{\|h\left(\mathbf{M}^{(i+1)} - \mathbf{M}^{(i)}\right)\|_2^2}$

to our minimisation problem as follows:

$$\underset{\mathbf{M}\in\mathcal{B}^{+}}{\text{minimise}}\ \frac{1}{2}\|\mathbf{Y} - h\left(\mathbf{M}\right)\|_2^2 + rG\left(\mathbf{M}\right), \tag{3.26}$$

where r is a constant that balances the data fidelity term with the regularisation term and g is a function incorporating prior information. Note that we can easily incorporate the SVD compression to this formulation by changing h to h_V and solving for $\widehat{\mathbf{M}}$.

We propose to solve problem (3.26) by computing at iteration i:

$$\mathbf{M}^{(i+1)} = \text{prox}_{r\mu^{(i)}G(\cdot)}\left(\mathcal{P}_{\mathcal{B}^{+}}\left(\mathbf{M}^{(i)} - \mu^{(i)}h^{\dagger}\left(h\left(\mathbf{M}^{(i)}\right) - \mathbf{Y}\right)\right)\right), \tag{3.27}$$

We can further incorporate the adaptive dictionary to the reconstruction process. We dubbed this method regularised Adaptive-BLIP (rABLIP). The details can be seen in Algorithm 5. Note that the proximity step is performed after the line search. The line search is used to avoid local minimum in the minimisation problem. If we perform the proximity step as suggested by the forward-backward algorithm (before the projection step), the performance of the line search appears to be reduced.

Algorithm 5 Regularised Adaptive-BLIP

Input: $\mathbf{Y} \in \mathbb{C}^{Q\times L\times C}$, $\zeta < 1$, $\mathbf{M}^{(0)} \in \mathbb{C}^{N\times L}$
Iterations:
for $i = 0, 1, \ldots$ **do**
  $\mu = 2N/Q$, $\nu = 0$
  while $\mu > \nu$ **do**
    $\mu = \mu/2$
    Forward Step:
    $\overline{\mathbf{M}}^{(i)} = \mathbf{M}^{(i)} - \mu h^{\dagger}\left(h\left(\mathbf{M}^{(i)}\right) - \mathbf{Y}\right)$
    Dictionary Estimation:
    $\mathbf{\Phi^{(i+1)}} = \text{DictionaryEstimation}\left(\overline{\mathbf{M}}^{(i)}, \mathbf{\Phi}^{(i)}\right)$
    Projection Step:
    $\widetilde{\mathbf{M}}^{(i)} = \mathcal{P}_{\mathcal{B}^{+}\left(\mathbf{\Phi}^{(i+1)}\right)}\left(\widetilde{\mathbf{M}}^{(i)}\right)$
    Backtracking step:
    $\nu = \zeta \frac{\|\widetilde{\mathbf{M}}^{(i)} - \mathbf{M}^{(i)}\|_2^2}{\|h\left(\widetilde{\mathbf{M}}^{(i)} - \mathbf{M}^{(i)}\right)\|_2^2}$
  Proximity step:
  $\mathbf{M}^{(i+1)} = \text{prox}_{r\mu G(\cdot)}\left(\widetilde{\mathbf{M}}^{(i)}\right)$

In this manuscript, we will consider two versions of G.

$$G_{\ell_1}(\mathbf{M}) = \sum_{l=1}^{L} \|\mathbf{\Psi}\mathbf{M}_{:,l}\|_1, \tag{3.28}$$

where $\mathbf{\Psi}$ is the 2-D Haar Wavelet transform with four levels. This function promotes sparsity in the wavelet dictionary and will be referred as ℓ_1-rABLIP. The second considered function is

$$G_{\text{TV}}(\mathbf{M}) = \sum_{l=1}^{L} \|\mathbf{M}_{:,l}\|_{TV}, \tag{3.29}$$

where $\|\cdot\|_{TV}$ denotes the total variation norm. This function promotes sparsity of the gradient and will be referred as TV-rABLIP.

3.4 Results

In this section, we present the procedure used to evaluate the reconstruction with simulated data using a simulated phantom. We will consider the acquisition detailed in [32], with eight coils (i.e. $C = 8$) and the corresponding sensitivity maps can be seen in Fig. 3.5.

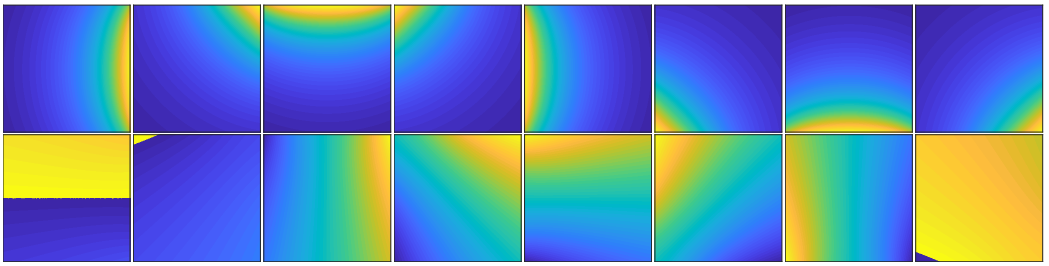


FIGURE 3.5: Coil sensitivity maps used in the simulated phantom experiments. First row corresponds to the magnitude and the second row corresponds to the phase.

We use the simulated phantom proposed in [33], with five tissues: adipose, WM, GM, muscle and cerebrospinal fluid (CSF). Two sampling patterns are considered, an Echo-planar Imaging (EPI) [34, 35] sampling used in [27]. and the spiral sampling used in [2]. In particular, the EPI sampling is performed using equally spaced lines of the k-space and the spiral sampling scheme uses a variable density spiral with 89 interleaves. The sampling pattern of the first acquisition instance $L = 1$ can be seen in Figure 3.6.

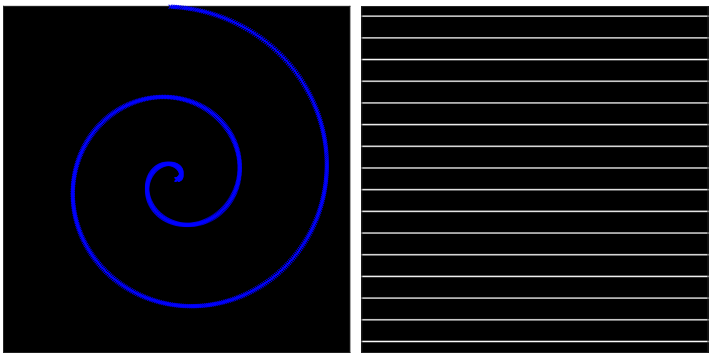


FIGURE 3.6: Sampling patterns for $L = 1$. (Left) Spiral sampling with undersampling ratio $N/Q = 89.53$. (Right) EPI sampling with undersampling ratio $N/Q = 16$.

The EPG formalism is used for the non-linear mapping, with the flip angles α and repetition times TR as described in [32] and can be seen in Fig. 3.7. We investigate

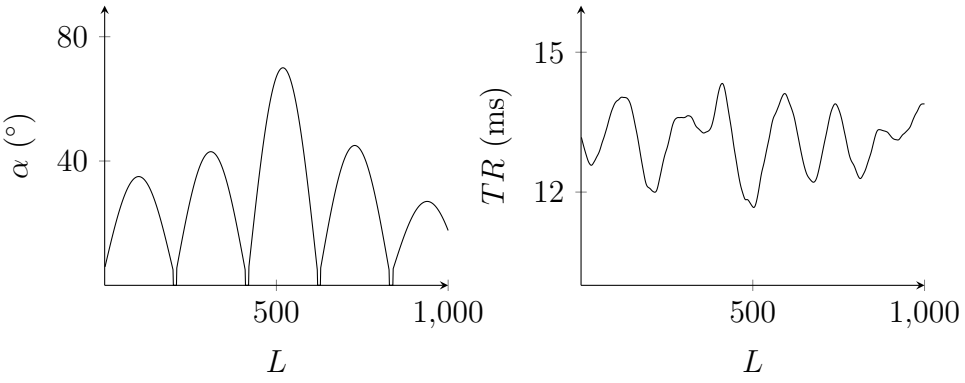


FIGURE 3.7: Acquisition parameters used in the simulated phantom experiments. (Left) Flip angles. (Right) Repetition times.

the effect of the measurement noise by varying the input SNR (iSNR in dB), defined as:

$$\text{iSNR} = 20 \log \left(\frac{\|h(\mathbf{M})\|_2}{(\sqrt{QLC}\sigma_{\mathbf{Y}}} \right), \tag{3.30}$$

where $\sigma_{\mathbf{Y}}$ is the standard deviation of the noise. We vary the iSNR from 10dB to 50dB. In addition, all the algorithms are evaluated using the SVD temporal compression of the sequence. The eigenvectors are computed using a densely sampled normalised dictionary taking all possible combinations of $T_1 \in [100, 5200]$ms in steps of 10ms and $T_2 \in [20, 600]$ms in steps of 2ms. The highest 100 eigenvalues of the dictionary can be seen in Fig. 3.8.

As seen in Fig. 3.8, a most of the energy is encapsulated into the first 10 eigenvalues. The algorithms are evaluated using a compression with 10 and 30 eigenvectors.

All the iterative algorithms are stopped when the following stopping criterion is satisfied:

$$\frac{|E^{(i+1)} - E^{(i)}|}{E^{(i+1)}} < 10^{-4}, \tag{3.31}$$

where $E^{(i)}$ is the value of the objective function at iteration i. We will refer as matching filter (MF) to the algorithm described in [2] for the full operator h and the method described in [26] for the temporal compression operator h_V. We will refer as iterative reconstruction (IR) to the algorithm described [27] for the full operator and

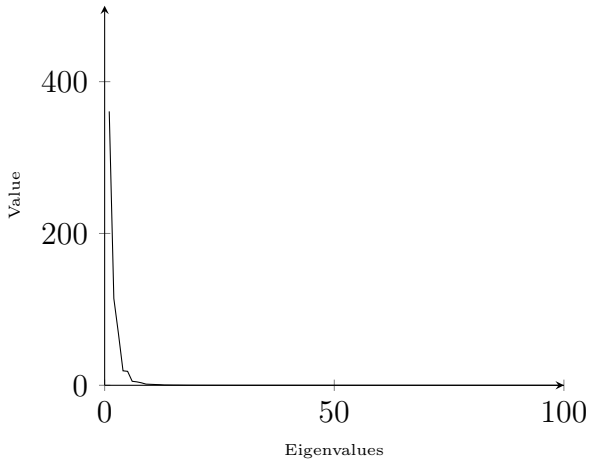


FIGURE 3.8: Highest 100 eigenvalues of the densely sampled normalised dictionary.

the method described [28] for the temporal compression operator h_V. The method described [28] with the low-pass filter will be referred as the iterative reconstruction with low-pass filter (IR + lp). Note that the IR + lp is just used to reconstruct the spiral sampling simulations because the in the EPI sampling simulations all frequencies are sampled. For the adaptive algorithms, the number of clusters is set to $K = 20$. For the regularised algorithms, we tuned r for the highest iSNR scenario and we keep it fixed for all other scenarios.

The results of all algorithms can be seen in Figures 3.9-3.11. As seen in the results, the reconstructions using the temporal compression not only have similar reconstruction than the full operator but in most cases are slightly better. Note that since the noise is not correlated with the temporal compression part of the noise is filtered. We can see that there is a significant improvement from the MF algorithms to the IR algorithms for simulations with iSNR$>$ 30dB. As mentioned before MF algorithms can be seen as the first iteration of the IR algorithms this explain the improvement. Note that for all the iterative algorithms in iSNR$<$ 30dB scenarios with the spiral sampling, the T_2 reconstructions are similar or worst than MF algorithms. This is due to noise over-fitting. IR+lp has a similar reconstructions due since at each iteration a low-pass filter is applied to the reconstructions removing part of the noise.

The proposed algorithms outperform the IR and MF algorithms. This is due to the ability to refine the dictionary through the iterations. Note that regularised algorithms

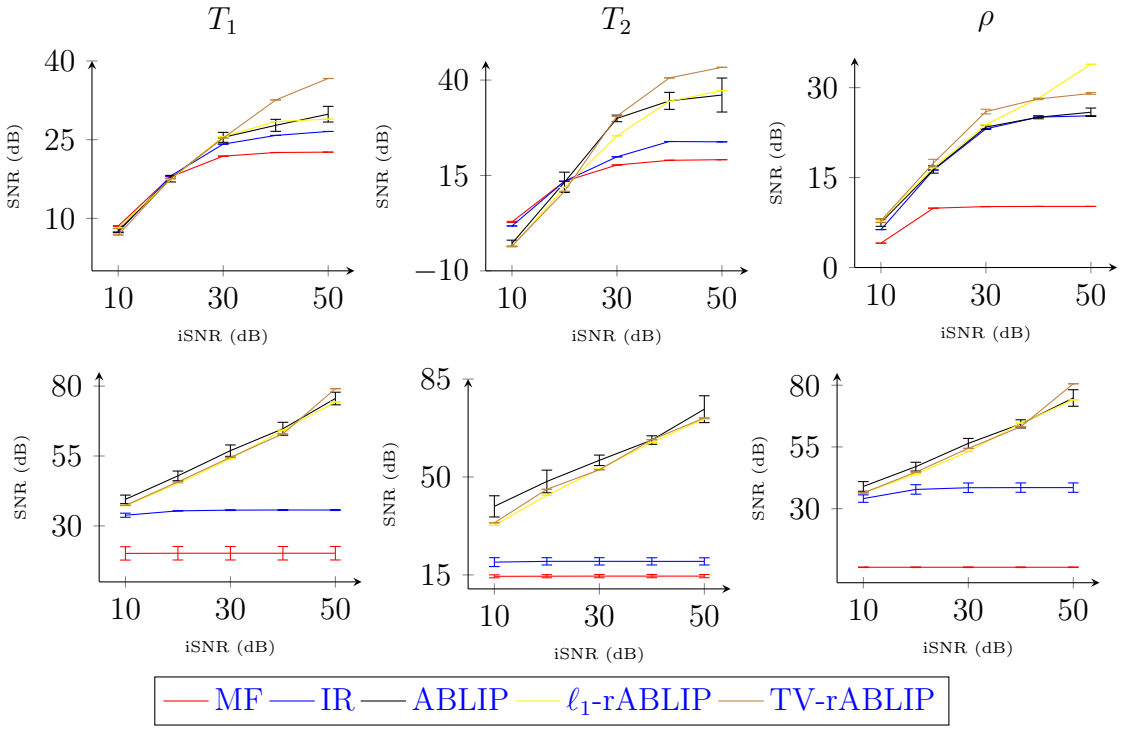


FIGURE 3.9: Results of the algorithms using the full acquisition operator h. The first row corresponds to the simulations with spiral sampling and the second row corresponds to the EPI sampling.

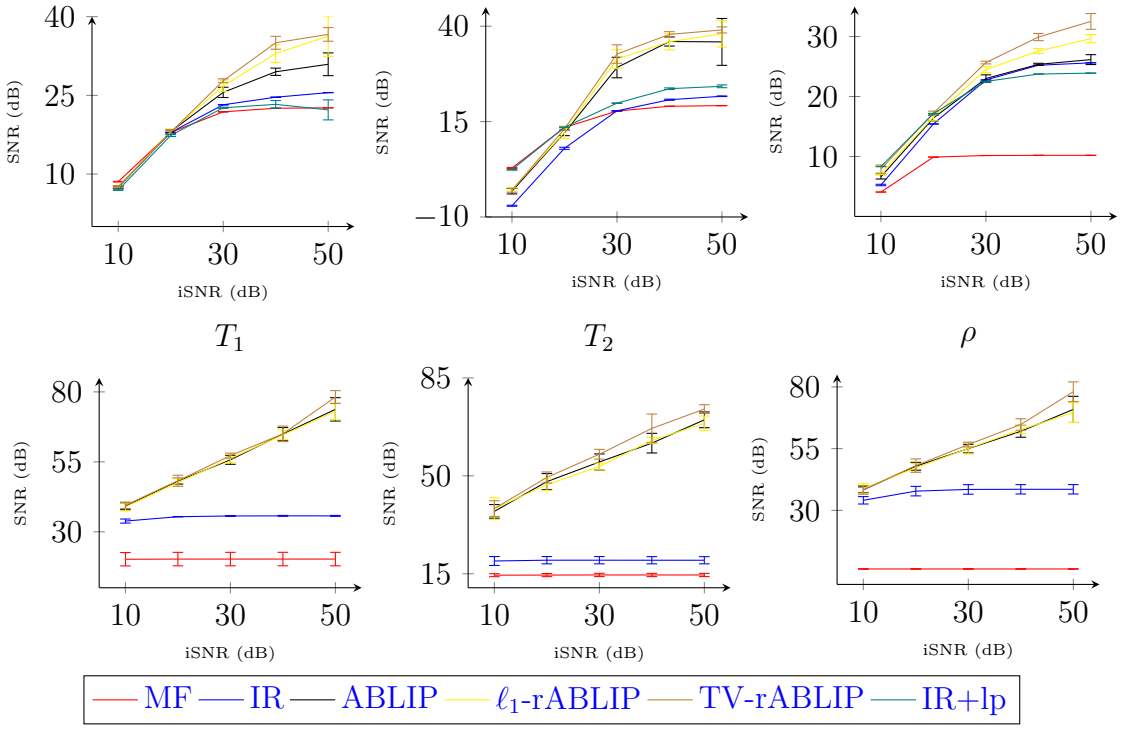


FIGURE 3.10: Results of the algorithms using the temporal compression operator with 10 eigenvectors h_V. The first row corresponds to the simulations with spiral sampling and the second row corresponds to the EPI sampling.

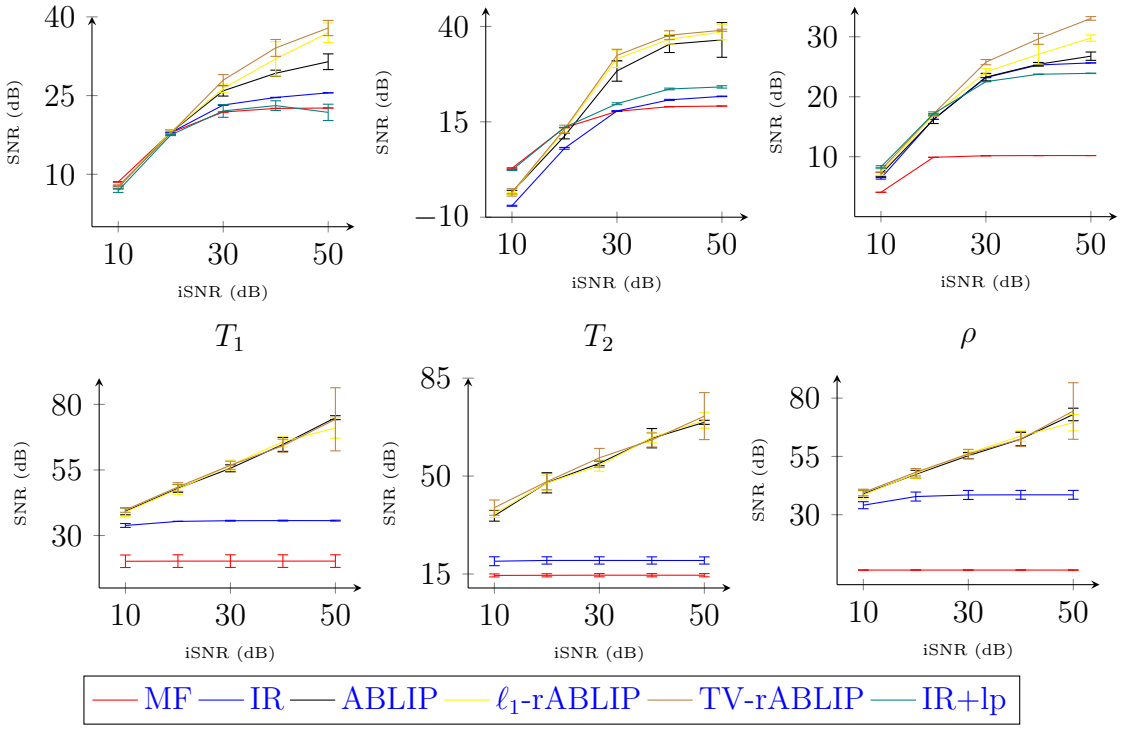


FIGURE 3.11: Results of the algorithms using the temporal compression operator with 30 eigenvectors h_V. The first row corresponds to the simulations with spiral sampling and the second row corresponds to the EPI sampling.

with the EPI sampling perform similarly than ABLIP but significantly improve the reconstructions in the case of the spiral sampling. This is improvement is explained by the CS theory [6, 12, 13, 36]. Note that the G_{TV} regularisation outperforms G_{ℓ_1}. This is due to the TV-norm ability to promote smooth images while preserving the edges.

In Figures 3.12-3.14, an example of the reconstructions of each algorithm can be seen for iSNR= 50dB scenario with spiral sampling. Note that the IR and ABLIP reconstructions presents high frequency artefacts while the regularised algorithms the artefacts are significantly attenuated.

In Figures 3.15-3.17, an example of the reconstructions of each algorithm can be seen for iSNR= 50dB scenario with EPI sampling. Note that with this sampling the IR, ABLIP and the regularised algorithms perform similarly. The main reason is that we have enough high frequency information in the measurements and adding

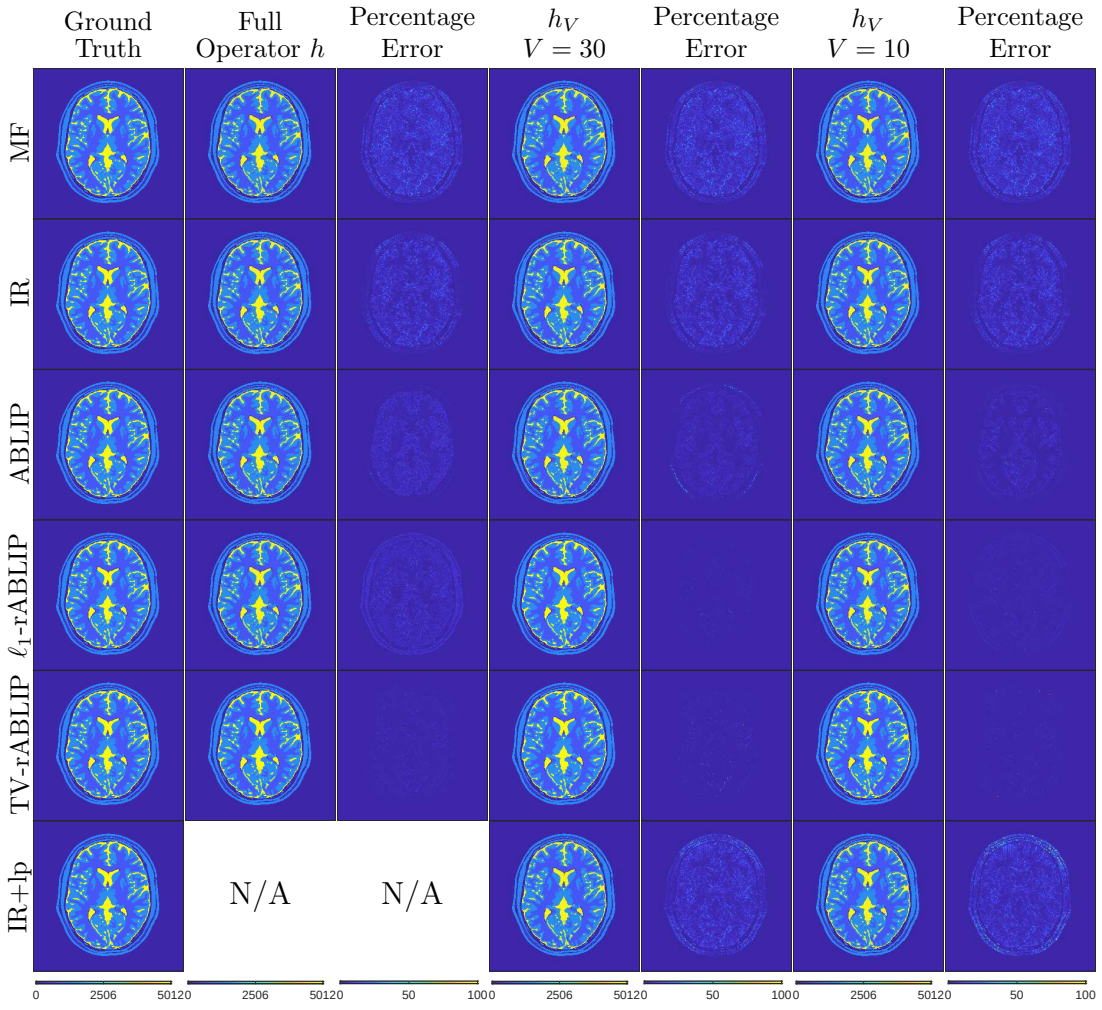


FIGURE 3.12: Example of the T_1 reconstructions with iSNR of 50dB of the spiral sampling. From first to last column: Ground truth images, reconstructions using h, reconstruction percentage error, reconstructions using h_V with 10 eigenvectors, reconstruction percentage error, reconstructions using h_V with 30 eigenvectors, reconstruction percentage error. From first to last row: MF, IR, ABLIP, ℓ_1-rABLIP, TV-rABLIP and IR+lp.

spatial regularisation only improve slightly the reconstructions and in some cases the performance is even reduced.

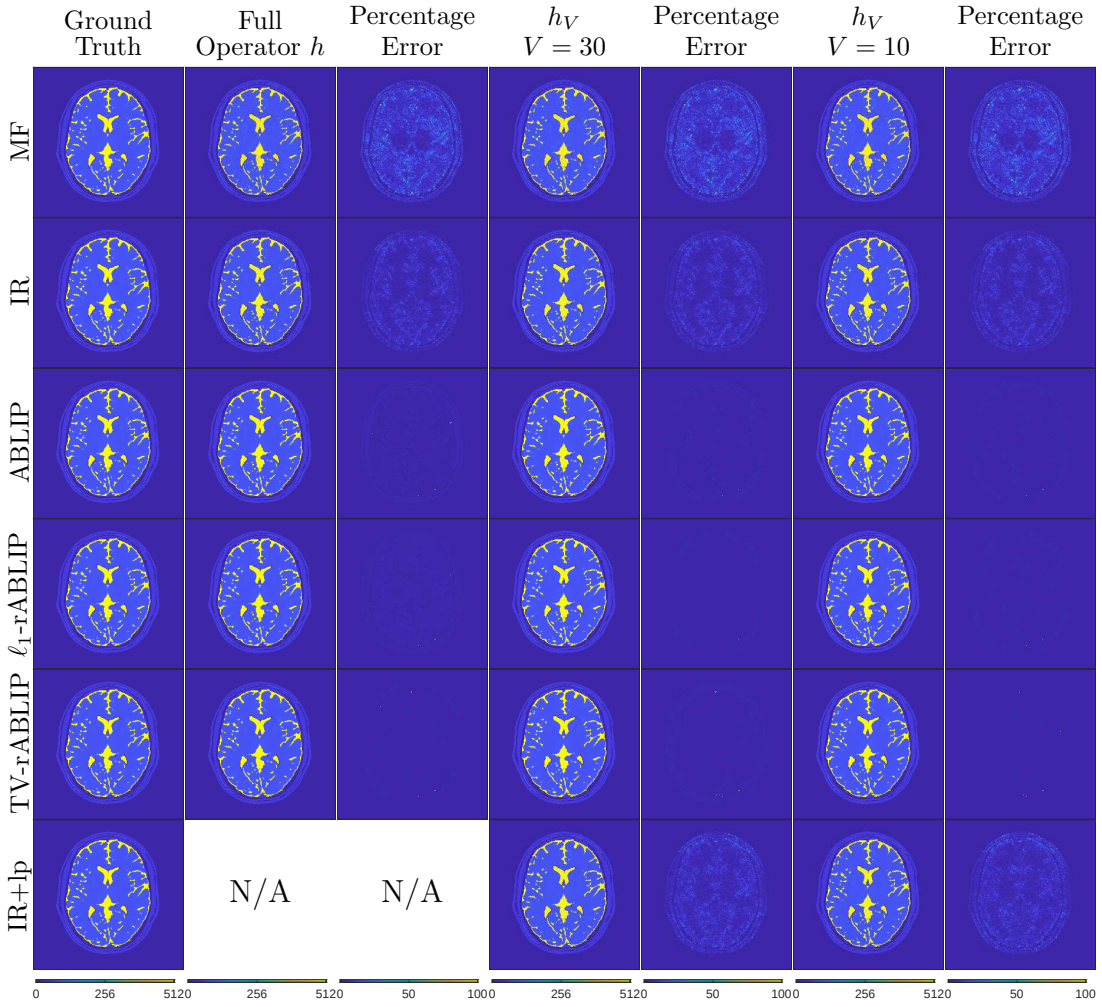


FIGURE 3.13: Example of the T_2 reconstructions with iSNR of 50dB of the spiral sampling. From first to last column: Ground truth images, reconstructions using h, reconstruction percentage error, reconstructions using h_V with 10 eigenvectors, reconstruction percentage error, reconstructions using h_V with 30 eigenvectors, reconstruction percentage error. From first to last row: MF, IR, ABLIP, ℓ_1-rABLIP, TV-rABLIP and IR+lp.

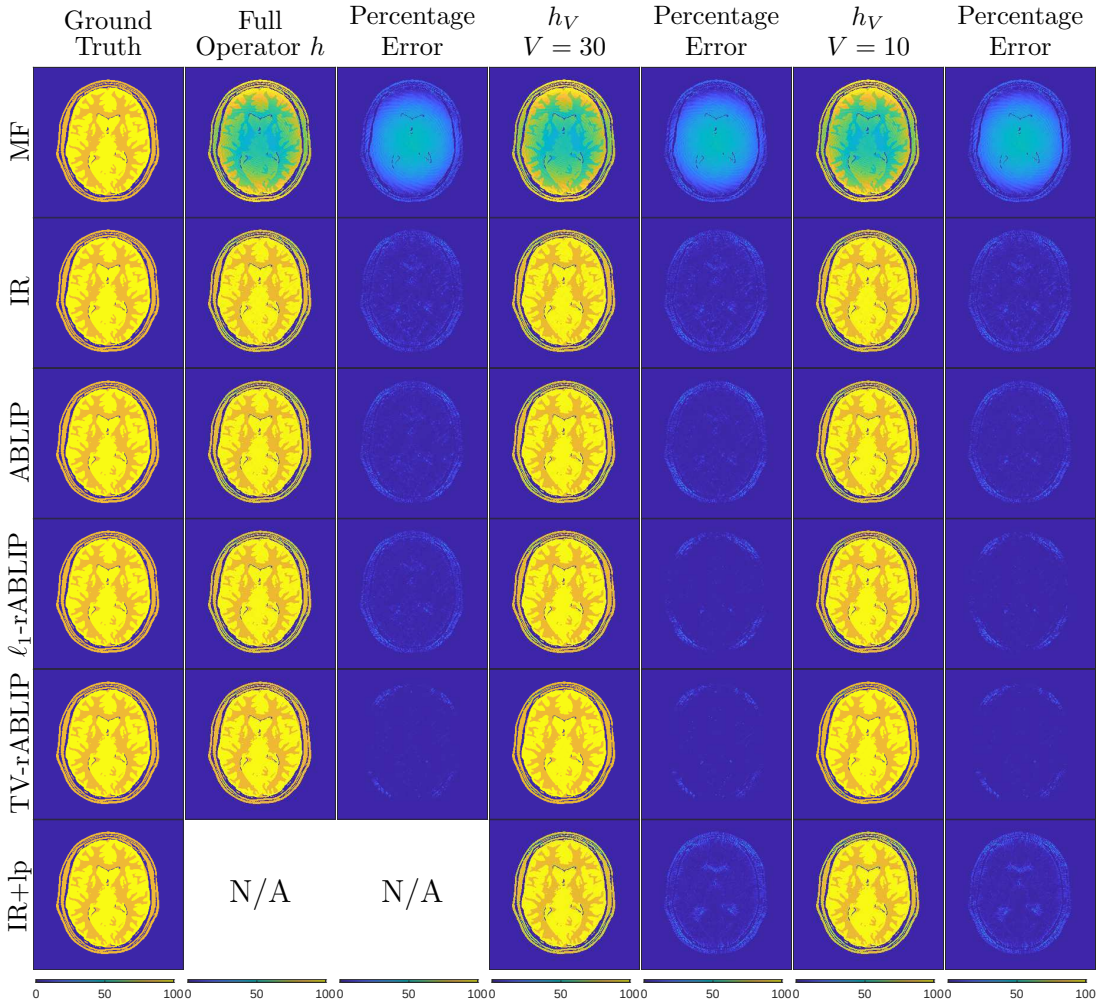


FIGURE 3.14: Example of the ρ reconstructions with iSNR of 50dB of the spiral sampling. From first to last column: Ground truth images, reconstructions using h, reconstruction percentage error, reconstructions using h_V with 10 eigenvectors, reconstruction percentage error, reconstructions using h_V with 30 eigenvectors, reconstruction percentage error. From first to last row: MF, IR, ABLIP, ℓ_1-rABLIP, TV-rABLIP and IR+lp.

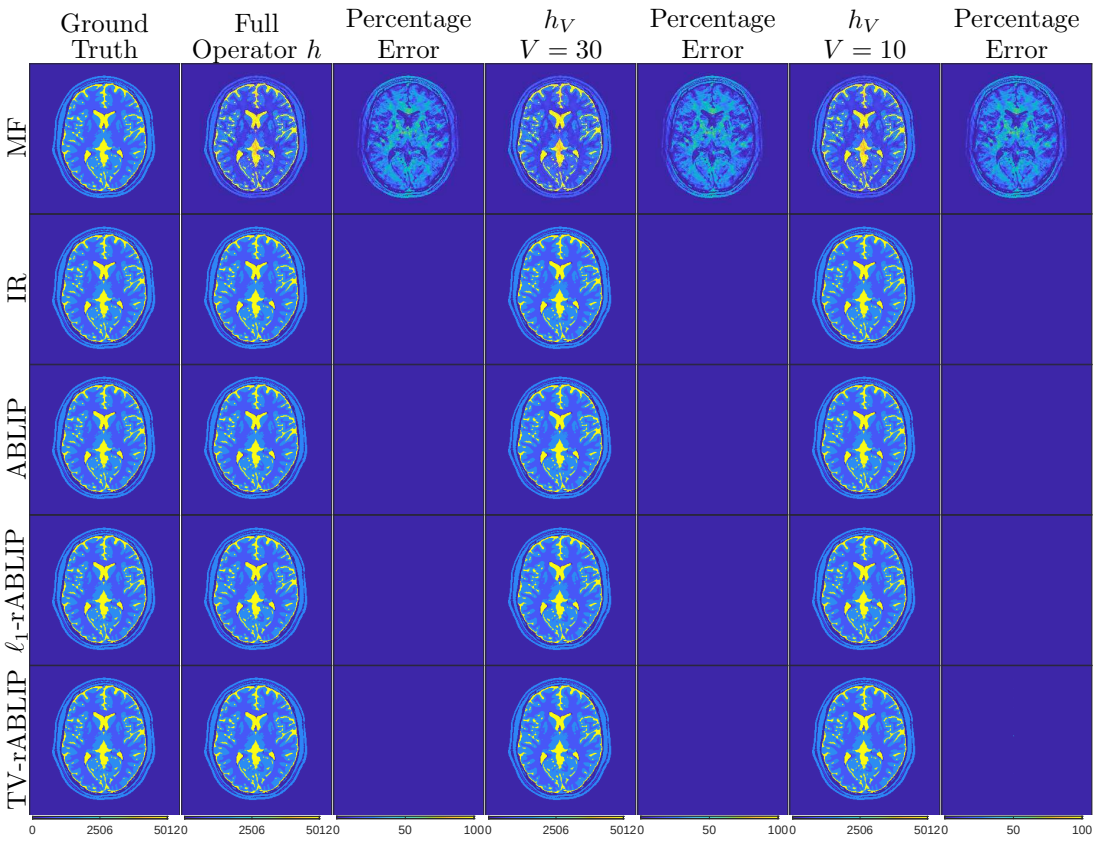


FIGURE 3.15: Example of the T_1 reconstructions with iSNR of 50dB of the EPI sampling. From first to last column: Ground truth images, reconstructions using h, reconstruction percentage error, reconstructions using h_V with 10 eigenvectors, reconstruction percentage error, reconstructions using h_V with 30 eigenvectors, reconstruction percentage error. From first to last row: MF, IR, ABLIP, ℓ_1-rABLIP and TV-rABLIP.

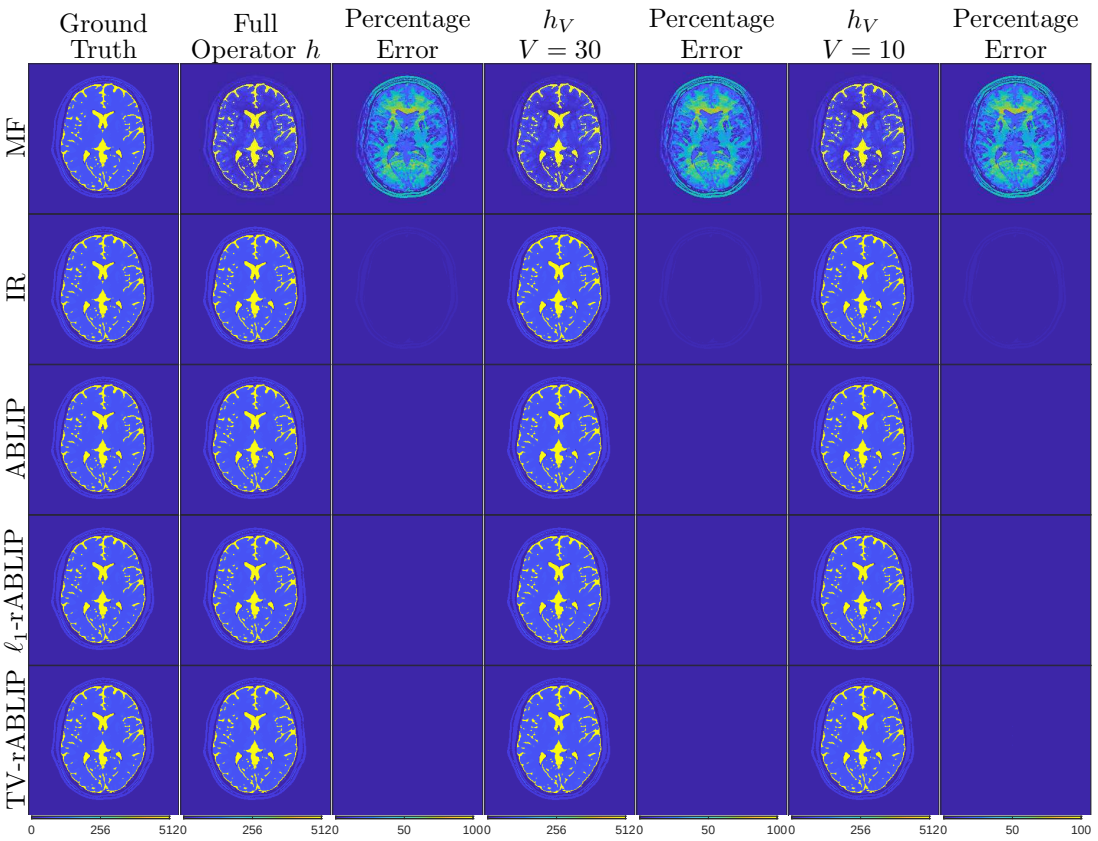


FIGURE 3.16: Example of the T_2 reconstructions with iSNR of 50dB of the EPI sampling. From first to last column: Ground truth images, reconstructions using h, reconstruction percentage error, reconstructions using h_V with 10 eigenvectors, reconstruction percentage error, reconstructions using h_V with 30 eigenvectors, reconstruction percentage error. From first to last row: MF, IR, ABLIP, ℓ_1-rABLIP and TV-rABLIP.

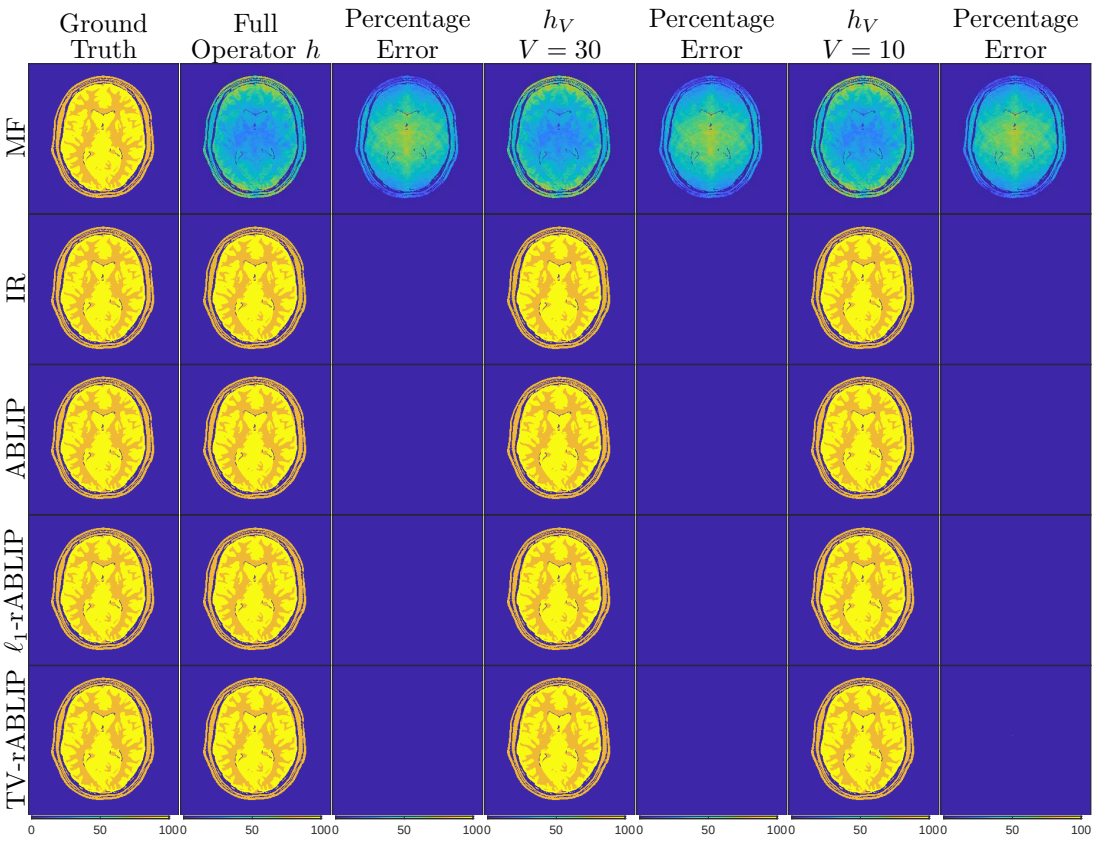


FIGURE 3.17: Example of the ρ reconstructions with iSNR of 50dB of the EPI sampling. From first to last column: Ground truth images, reconstructions using h, reconstruction percentage error, reconstructions using h_V with 10 eigenvectors, reconstruction percentage error, reconstructions using h_V with 30 eigenvectors, reconstruction percentage error. From first to last row: MF, IR, ABLIP, ℓ_1-rABLIP and TV-rABLIP.

Chapter 4

Partial Volume in MRF

In general, qMRI techniques, particularly MRF-based methods [24, 26, 27, 37–40], assume that a voxel contains at most one type of tissue, e.g. white matter (WM), grey matter (GM), etc. This assumption is not suitable in practice. Consequently, voxels containing multiple tissue types may be assigned with incorrect parameters. This problem is known as the Partial Volume Effect (PVE) and appears in all medical imaging modalities with limited spatial resolution [41]. An example of PVE is given in Fig. 4.1. The left image shows a spatial distribution of T_1 in a simulated brain. The right image shows a reconstruction using voxels four times bigger and assuming a single tissue per voxel. All low-resolution voxels at the edge between tissues contain partial volumes, which implies a wrong estimate (single wrong value of T_1 rather than multiple values).

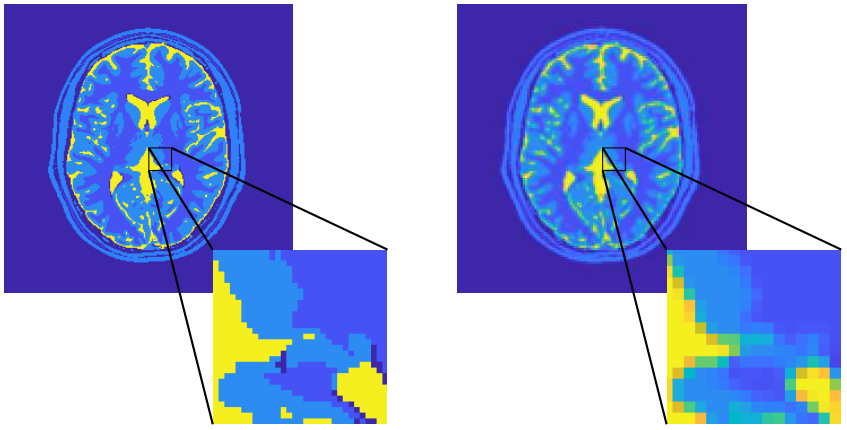


FIGURE 4.1: Partial volume effect in a T_1 parameter map. Left: true T_1 parameter map. Right: low resolution reconstruction.

The PVE has been analysed in the supplementary material of [2]. In this work, using a least-squares method, the signal is decomposed as a weighted sum of at most three distinct signals, each representing a different tissue. Although this method was shown to be robust to noise for long sequences, since it necessitates both information about

the spatial distribution of the PV voxels and the true components of the original signal (which are unknown in practice), it is not adapted to handle *in vivo* data.

An extension of this approach has been proposed in [42], where the tissue parameters are learnt using a clustering approach on the parameter maps, obtained by the matched filter. Then the data is matched with a PV dictionary varying the tissue proportion. This method has shown good results when considering long sequences, but the reconstruction quality is limited by the precision of the dictionary elements obtained by the MRF solution and the precision of the tissue proportions used to generate the PV dictionary. Moreover, the size of the PV dictionary increases exponentially with the maximum number of tissues allowed in a voxel and, the parameters for the PV dictionary are manually selected from the clustering results. In addition, all the reconstructions are performed with high aliased images requiring more acquisitions for accurate results. Additionally, for short sequences, the noise in the measurements and the sampling of the manifold of fingerprints describing the signal can significantly affect the estimations.

More recently, a Bayesian method was proposed in [43], to tackle the PVE in MRF (we will refer to this method as Bayesian-MRF). The authors show that their approach estimates the parameters of the PV voxels. However, due to the high aliasing effect encountered with undersampled noisy data, this estimation comes at the cost of an increased acquisition time with respect to traditional MRF based reconstructions (*i.e.* three times longer sequences than traditional MRF). Furthermore, to obtain accurate results, this method relies on a high sampling of the fingerprint manifold, resulting in a high computational cost (in terms of both reconstruction time and memory requirement). While this method was formulated as a convex optimisation problem in the Bayesian framework, the algorithm in [43] removes dictionary entries at each iteration and consequently does not ensure that the algorithm converges to the maximum *a posteriori* estimate.

In the first part of the chapter, the model considering partial volumes is introduced and the challenges that arise are discussed. A description of the state-of-the-art techniques is presented in the second part. The last part of the chapter a novel PVE

model for MRF is presented. To solve the resulting non-convex constrained minimisation problem, a greedy approximate projected gradient descent method, dubbed GAP-MRF, is developed. It can be seen as a generalisation of BLIP method for PVE. It consists of a projected gradient descent algorithm, where the projection is computed inexactly, through a memory-efficient greedy approach. The proposed method is also generalised to compensate for phase errors in the model, due to timing or coil sensitivity errors. This compensation is performed using an alternating minimisation approach [3, 44–47].

4.1 Partial Volume Model

The model described in the previous chapter considers that each voxel contains at most one element. PV voxels are introduced due to the spatial discretisation in the acquisition process. The magnetisation sequence can be described as $\mathbf{M} = \mathbf{X}\mathbf{\Phi}$, where $\mathbf{X} \in \mathbb{R}_+^{N\times D}$ is a sparse mixing matrix (each line of $\mathbf{X}$ represents the proton densities associated with a specific voxel, and would contain more than a nonzero value only for voxels with partial volumes), and $\mathbf{\Phi} \in \mathbb{C}^{D\times L}$ is the over-complete dictionary of fingerprints, introduced in [2], as a discrete sampling of the low dimensional manifold B. $\mathbf{\Phi}$ is constructed from D samples of $\mathcal{M}$, stored in a matrix $\boldsymbol{\theta} \in \mathbb{R}^{D\times P}$. Due to the smoothness of B, $\mathbf{\Phi}$ is highly coherent. Consequently, the estimation of $\mathbf{X}$ from highly undersampled noisy data is expected to fail without additional priors. Leveraging CS theory [6, 12, 13, 36], the sparsest matrix $\mathbf{X}$, fitting the measurement model, can be found by solving:

$$\underset{\mathbf{X}\in\mathbb{R}_+^{N\times D}}{\text{minimise}} \ \|\mathbf{X}\|_0 \text{ subject to } \|\mathbf{Y} - h(\mathbf{X}\mathbf{\Phi})\|_2 \leq \epsilon, \tag{4.1}$$

where $\epsilon > 0$ is a bound chosen according to the acquisition noise level. Since this function is non-convex and non-differentiable, problem (4.1) is difficult to solve in practice, in particular in the context of high dimensional problems (usually, $D \sim 10^6$ and $L \sim 10^3$). Thus, the non-convexity of the ℓ_0 pseudo-norm is often relaxed by

the use of the ℓ_1-norm [48]. Nevertheless, $\mathbf{\Phi}$ being highly coherent, this convex relaxation cannot be used to correctly estimate the coefficients of $\mathbf{X}$ [49]. We presented a manuscript in the Signal and Processing with Adaptive Sparse Structured Representation (SPARS) Workshop in 2017. This manuscript shows the need of a post-processing step after the solving the convex relaxation problem:

$$\underset{\mathbf{X}\in\mathbb{R}_{+}^{N\times D}}{\text{minimise}} \|\mathbf{X}\|_1 \text{ subject to } \|\mathbf{Y} - h(\mathbf{X}\mathbf{\Phi})\|_2 \leq \epsilon. \tag{4.2}$$

4.2 MRF Techniques with Partial Volumes

In this section, the state-of-the-art methods for MRF with partial volume are presented. In particular, the partial volume dictionary proposed in [42], convex optimisation method proposed in [50] and the Bayesian formulation in [43].

4.2.1 Partial Volume Dictionaries

In [42] a method is proposed to estimate the tissue proportion from MRF acquisitions. This method uses the K-means algorithm [51] to cluster the parameters obtained by the matching filter proposed in [2], the obtained clusters are manually selected to represent the main tissues in the volume. The idea behind the clustering is to find the parameters of the pure voxels (voxels containing only one kind of tissue), and create a partial volume dictionary as a combination of the pure elements. The partial volume dictionary is constructed by varying the proportion of the dictionary elements of the selected clusters. Finally, a matched filter is performed using this dictionary to estimate the tissue proportion of the voxels. Let $\mathbf{\Phi}_c \in \mathbb{C}^{T\times L}$ be the dictionary of the selected clusters, then the partial volume dictionary is constructed by:

$$\mathbf{\Phi}_{PV} = \mathbf{P}\mathbf{\Phi}_c \tag{4.3}$$

where the rows $\mathbf{P} \in \mathbb{R}^{P \times T}$ correspond to all the feasible proportions of the elements in $\mathbf{\Phi}_c$. Note that the size of $\mathbf{\Phi}_{PV}$ grows exponentially depending on the maximum number of tissues per voxel.

4.2.2 Regularised Optimisation

Based on one of the recent methods in sparse recovery, a re-weighted ℓ_1-norm method has been proposed for MRF with partial volumes [50]. This method aims to solve the sequence of convex problems:

$$\begin{aligned} \underset{\widehat{\mathbf{M}} \in \mathbb{C}^{N \times V}, \mathbf{X} \in \mathbb{R}_+^{N \times D}}{\operatorname{argmin}} \quad & \left\| \mathbf{Y} - h_V\left(\widehat{\mathbf{M}}\right) \right\|_2^2 + r\left(\mathbf{X}\right) \\ \text{subject to} \quad & \widehat{\mathbf{M}} = \mathbf{X}\widehat{\mathbf{\Phi}} \end{aligned} \tag{4.4}$$

where r is a regularisation function promoting sparsity. The authors of [50] proposed to solve (4.4) using a method based on the Alternating Direction Method of Multipliers (ADMM). In particular, the augmented Lagrangian for (4.4) is:

$$\mathcal{L}_\gamma\left(\widehat{\mathbf{M}}, \mathbf{X}, \mathbf{L}\right) = \left\| \mathbf{Y} - h_V\left(\widehat{\mathbf{M}}\right) \right\|_2^2 + r\left(\mathbf{X}\right) + \left\langle \mathbf{L}, \widehat{\mathbf{M}} - \mathbf{X}\widehat{\mathbf{\Phi}} \right\rangle + \frac{\gamma}{2} \left\| \widehat{\mathbf{M}} - \mathbf{X}\widehat{\mathbf{\Phi}} \right\|_2^2 \tag{4.5}$$

where $\mathbf{L} \in \mathbb{C}^{N \times V}$ is the dual variable. While it is not clear what is the form of r, the authors of [50] update the variable $\mathbf{X}$ as the solution of the re-weighted ℓ_1-norm by solving a sequence of J problems of the form:

$$\underset{\mathbf{X} \in \mathbb{R}_+^{N \times D}}{\operatorname{argmin}} \quad \frac{1}{2} \left\| \mathbf{X}\widehat{\mathbf{\Phi}} - \widehat{\mathbf{M}} \right\|_2^2 + \left\| \mathbf{W}^{(j)} \circ \mathbf{X} \right\|_1 \tag{4.6}$$

where $j \in \{1, 2, \ldots, J\}$ is the index of the problem, $\circ$ is the Hadamard product and $\mathbf{W}^{(j)} \in \mathbb{R}_+^{N \times D}$ is the matrix of weights, initially a matrix with all its elements equal to 1. The weights are updated after solving each convex problem as proposed in [52]:

$$\mathbf{W}_{n,d}^{(j+1)} = \left(\widehat{\mathbf{\Phi}}_{d,:} \left(\epsilon \mathbf{I}^d + \widehat{\mathbf{\Phi}}^\dagger \operatorname{Diag}\left(\mathbf{W}_{n,:}^{(j)}\right)^{-1} \operatorname{Diag}(\mathbf{X}_{n,:}) \widehat{\mathbf{\Phi}} \right)^{-1} \left[\widehat{\mathbf{\Phi}}_{d,:} \right]^\dagger \right)^{1/2}, \tag{4.7}$$

where $\epsilon > 0$ is a small value compared to the values of $\mathbf{X}_{n,:}$ to avoid dividing by 0, $\mathbf{I}^d$ is the identity matrix of size $d \times d$. On the opposite to other choices of re-weighting such as [53], the correlation of the dictionary is included in the update to avoid inaccuracies in the initial estimation. As stated in [50], intuitively this promotes sparsity by penalising coefficients of $\mathbf{X}$ that does not contribute to the fit in previous solutions.

As shown in the results, the method is able to estimate the partial volumes only in high iSNR scenarios. This suggests that the prior is not able to resolve the partial volume due to the correlation of the dictionary elements. It is not clear if the algorithm converges to a solution, but the authors stop the algorithm after ten ADMM iterations.

4.2.3 Bayesian Inference

Methods based on Bayesian inference are used to provide a solution for inverse problems by analysing a *posterior* probability density distribution. This probability density is described by Bayes's formula which is in terms of the likelihood and prior probability density. In [43], a Bayesian method was proposed for partial volume estimations in the context of MRF. In this work, similarly to [2,26], a dirty version of the magnetisation sequence $\overline{\mathbf{M}}$ is estimated using (3.4). Since the noise assumed to be Gaussian with zero mean and constant variance $\sigma_{\mathbf{Y}}$, the authors propose a Gaussian likelihood density in a voxelwise fashion as:

$$\pi\left(\overline{\mathbf{M}}_{n,:}|\mathbf{X}_{n,:}\right) \propto \exp\left(-\frac{1}{2\sigma_{\mathbf{Y}}}\left\|\overline{\mathbf{M}}_{n,:} - \mathbf{X}_{n,:}\mathbf{\Phi}\right\|_2^2\right), \tag{4.8}$$

where $\propto$ denotes proportionality. The values of $\mathbf{X}_{n,:}$ are assumed to be independent and normally distributed with zero mean and variance $\boldsymbol{\theta}_d$. This *prior* is then described as:

$$\pi(\mathbf{X}_{n,:}) = \frac{\det\left(\text{Diag}\left(\boldsymbol{\theta}^{-1/2}\right)\right)}{2\pi} \exp\left(-\frac{1}{2}\left\|\mathbf{X}_{n,:}\text{Diag}\left(\boldsymbol{\theta}^{-1/2}\right)\right\|_2^2\right), \tag{4.9}$$

where det is the determinant. The covariance matrix $\text{Diag}\left(\boldsymbol{\theta}\right)$ is assumed to be independent and identically distributed, following a Gamma distribution, with shape and

scale parameterised by α and β, as follows:

$$\pi\left(\boldsymbol{\theta}_d\right) \propto \boldsymbol{\theta}_d^{\alpha-1} \exp\left(\frac{-\theta_d}{\beta}\right). \tag{4.10}$$

The *posterior* density distribution is obtained by applying Bayes's formula as:

$$\pi\left(\mathbf{X}_{n,:}, \boldsymbol{\theta} | \overline{\mathbf{M}}_{n,:}\right) \propto \pi\left(\overline{\mathbf{M}}_{n,:} | \mathbf{X}_{n,:}\right) \pi(\mathbf{X}_{n,:}) \pi\left(\boldsymbol{\theta}\right). \tag{4.11}$$

The maximum *a posteriori* (MAP) estimator is used to define the target minimisation function as:

$$\begin{aligned} \underset{\mathbf{X}_{n,:} \in \mathbb{C}^{1\times D}, \boldsymbol{\theta} \in \mathbb{R}^{D\times 1}}{\operatorname{argmin}} \quad & \frac{1}{\sigma^2} \left\| \overline{\mathbf{M}}_{n,:} - \mathbf{X}_{n,:} \boldsymbol{\Phi} \right\|_2^2 + \frac{1}{2} \left\| \mathbf{X}_{n,:} \mathrm{Diag}\left(\boldsymbol{\theta}\right)^{1/2} \right\|_2^2 + \\ & \frac{1}{\beta} \sum_{d=1}^{D} \boldsymbol{\theta}_d + \left(\alpha + \frac{3}{2}\right) \sum_{d=1}^{D} \log \boldsymbol{\theta}_d. \end{aligned} \tag{4.12}$$

The authors solve the problem (4.12), by alternating:

$$\boldsymbol{\theta}_d^{(i)} = \beta \left(\eta + \left(\frac{\left(\mathbf{X}_{n,d}\right)^2}{2\beta} + \eta^2 \right)^{1/2} \right), \tag{4.13}$$

where $\eta = (\alpha - 3/2)/2$, and

$$\mathbf{X}_{n,:}^{(i)} = \mathrm{Diag}(\theta)^{1/2} \mathbf{w}^{(i)}, \tag{4.14}$$

where $\mathbf{w}^{(i)}$ is obtained by minimising:

$$\mathbf{w}^{(i)} = \underset{\mathbf{w} \in \mathbb{C}^{1\times D}}{\operatorname{argmin}} \quad \left\| \overline{\mathbf{M}}_{n,:} - \mathbf{w} \mathrm{Diag}\left(\boldsymbol{\theta}\right)^{1/2} \boldsymbol{\Phi} \right\|_2^2 + r \left\| \mathbf{w} \right\|_2^2 \tag{4.15}$$

where r is a regularisation parameter that balance the weight of the data fidelity term and the energy of the vector $\mathbf{w}$.

4.3 Greedy Approximate Projection with Partial Volumes

In this section, we present a novel partial volume problem for MRF and a Greedy algorithm to solve it. Part of this work was presented in the International Society for Magnetic Resonance in Medicine (ISMRM) Workshop on Magnetic Resonance Fingerprinting in 2017 and published in the Inverse Problems Journal - IOP Science in February 2020.

4.3.1 Proposed Model

To overcome these difficulties, similarly to the BLIP approach, we propose to

$$\underset{\mathbf{M}\in\mathcal{B}_{\mathcal{S}_+}(\mathbf{\Phi})}{\text{minimise}}\ \frac{1}{2}\|\mathbf{Y}-h(\mathbf{M})\|_2^2 \tag{4.16}$$

where

$$\mathcal{B}_{\mathcal{S}_+}(\mathbf{\Phi}) = \left\{\mathbf{M}\in\mathbb{C}^{N\times L}\,|\,\mathbf{M}=\mathbf{X}\mathbf{\Phi} \text{ with } \mathbf{X}\in\mathcal{S}_+\right\}, \tag{4.17}$$

$$\mathcal{S}_+ = \bigcap_{s=1}^{4}\mathcal{S}_s, \tag{4.18}$$

and, for every $s\in\{1,\ldots,4\}$, $\mathcal{S}_s$ is a closed non-empty subset of $\mathbb{R}^{N\times D}$ used to impose feasibility constraints on $\mathbf{X}$. These sets are defined below.

4.3.1.1 Positivity constraint

Since the proton densities of the imaged volume must be non-negative, we can restrict our solution to be in the positive orthant:

$$\mathcal{S}_1 = \mathbb{R}_+^{N\times D}. \tag{4.19}$$

4.3.1.2 Constraint on the number of tissues

Commonly MRF aims to obtain quantitative values of a small set of tissues. In practice, only $T \ll D$ elements of the dictionary $\mathbf{\Phi}$ are necessary to characterise $\mathbf{M}$. While T is unknown, we have a reasonable estimate for it. We propose to introduce a loose upper bound K, such that $T \leq K \leq D$, to limit the number of active dictionary elements. Let us define a set $\mathcal{D}_{\mathbf{X}}$ that is formed by the column indices of $\mathbf{X}$ with non-zero coefficients. To avoid noisy voxels, only rows with proton density greater than $\xi > 0$ (chosen according to the noise level) will be considered. Formally, this set is defined as $\mathcal{D}_{\mathbf{X}} = \{d \in \{1, \dots, D\} \,|\, (\exists n \in \mathcal{G}_{\mathbf{X}})\ \mathbf{X}_{n,d} \neq 0\}$, where $\mathcal{G}_{\mathbf{X}} = \{n \in \{1, \dots, N\} \,|\, \|\mathbf{X}_{n,:}\|_1 > \xi\}$. The set $\mathcal{D}_{\mathbf{X}}$ indicates the columns of $\mathbf{X}$ contributing to the magnetisation sequence. We can limit the number of used elements of the dictionary by upper bounding the cardinality of this set by K:

$$\mathcal{S}_2 = \left\{\mathbf{X} \in \mathbb{R}^{N \times D} \,|\, \mathrm{Card}(\mathcal{D}_{\mathbf{X}}) \leq K\right\}. \tag{4.20}$$

4.3.1.3 Constraint on the manifold neighbourhoods

The tissues of interest are unique and need to be sufficiently different to be distinguished. To incorporate this prior information in the reconstruction process, we define the neighbour set associated with each element $d \in \{1, \dots, D\}$ of the dictionary as:

$$\begin{aligned} \mathcal{N}_{\upsilon}(d) = \{d' \in \{1, \dots, D\} \backslash \{d\} \,|\, \\ (\forall p = \{1, \dots, P\})\, |\boldsymbol{\theta}_{d',p} - \boldsymbol{\theta}_{d,p}| < \upsilon \boldsymbol{\theta}_{d,p}\}, \end{aligned} \tag{4.21}$$

where $\upsilon > 0$. We define a set of all possible $\mathbf{X}$ such that, the parameters of each element in $\mathcal{D}_{\mathbf{X}}$ are sufficiently far from each other. Precisely, we constrict all the neighbour columns of each element in $\mathcal{D}_{\mathbf{X}}$ to be the null element $\mathbf{0}$ of $\mathbb{R}^N$:

$$\mathcal{S}_3 = \left\{\mathbf{X} \in \mathbb{R}^{N \times D} \,|\, \left(\forall d' \in \bigcup_{d \in \mathcal{D}_{\mathbf{X}}} \mathcal{N}_{\upsilon}(d)\right) \mathbf{X}_{:,d'} = \mathbf{0}\right\} \tag{4.22}$$

4.3.1.4 Constraint on the pure voxels

Due to the additive noise in model (3.1), some elements of $\mathbf{X}$ corresponding to non-used dictionary elements take non-zero values. In order to avoid these noisy elements in the reconstructions, we impose that at least $\kappa > 0$ rows (i.e. voxels) of $\mathbf{X}$ contain only one non-zero value for each active column of $\mathbf{X}$. These rows identify the *pure voxels*. This constraint can be formulated as follows:

$$\mathcal{S}_4 = \{\mathbf{X} \in \mathbb{R}^{N \times D} \mid (\forall d \in \mathcal{D}_{\mathbf{X}}) \parallel (\mathbf{X}_{n,d})_{n \in \mathcal{V}_{\mathbf{X}}} \parallel_0 \geq \kappa\} \tag{4.23}$$

where $\mathcal{V}_{\mathbf{X}} = \{n \in \{1, \ldots, N\} \mid \|\mathbf{X}_{n,:}\|_0 = 1\}$.

4.3.2 Proposed Algorithm

To solve problem (4.16), we use an iterative projected gradient descent method [54]. At each iteration $i \in \mathbb{N}$, this method updates $\mathbf{M}^{(i+1)}$ by computing a gradient step followed by a projection step:

$$\mathbf{M}^{(i+1)} = \mathcal{P}_{\mathcal{B}_{S_+}(\mathbf{\Phi})} \left(\mathbf{M}^{(i)} - \mu h^{\dagger} \big(h(\mathbf{M}^{(i)}) - \mathbf{Y} \big) \right), \tag{4.24}$$

where $\mu > 0$. In [27], it is shown that choosing $\mu \approx N/Q$ is theoretically justifiable. However, in order to ensure the stability of the iterative projected gradient descent algorithm and accelerate convergence, in [27, 55] the authors proposed to choose μ using a backtracking method. In order to handle efficiently the constraint $\mathcal{B}_{S_+}(\mathbf{\Phi})$, we propose to compute inexactly the projection onto this set in (4.24). The resulting method, named Greedy Approximate Projection for MRF (GAP-MRF), is described in Algorithm 6. It can be noticed that the GAP-MRF method and BLIP are solving similar problems, using the same algorithmic structure. In this context, as in [27], a condition on both L and the undersampling ratio N/Q might be derived for recovery guarantee. However, the investigation of such condition is beyond the scope of this manuscript.

Algorithm 6 GAP-MRF global iterations

Input: $\mathbf{Y} \in \mathbb{C}^{Q\times L\times C}$, $\zeta < 1$, $\mathbf{M}^{(0)} \in \mathbb{C}^{N\times L}$
Iterations:
for $i = 0, 1, \ldots$ **do**
  $\mu = 2N/Q$, $\nu = 0$
  while $\mu > \nu$ **do**
    $\mu = \mu/2$
    Gradient Step:
    $\overline{\mathbf{M}}^{(i)} = \mathbf{M}^{(i)} - \mu h^{\dagger}\left(h\left(\mathbf{M}^{(i)}\right) - \mathbf{Y}\right)$
    Projection Step:
    $\mathbf{M}^{(i+1)} \approx \mathcal{P}_{\mathcal{B}_{\mathcal{S}_+}(\mathbf{\Phi})}\left(\overline{\mathbf{M}}^{(i)}\right)$
    Backtracking step
    $\nu = \zeta \frac{\|\mathbf{M}^{(i+1)} - \mathbf{M}^{(i)}\|_2^2}{\|h\left(\mathbf{M}^{(i+1)} - \mathbf{M}^{(i)}\right)\|_2^2}$

4.3.3 Approximate projection

For every $\overline{\mathbf{M}} \in \mathbb{C}^{N\times L}$, we have:

$$\begin{aligned}\mathcal{P}_{\mathcal{B}_{\mathcal{S}_+}(\mathbf{\Phi})}\left(\overline{\mathbf{M}}\right) &= \underset{\mathbf{M}\in\mathcal{B}_{\mathcal{S}_+}(\mathbf{\Phi})}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{M} - \overline{\mathbf{M}}\|_2^2 \\ &= \underset{\mathbf{M}=\mathbf{X}\mathbf{\Phi},\, \mathbf{X}\in\mathcal{S}_+}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{X}\mathbf{\Phi} - \overline{\mathbf{M}}\|_2^2 \\ &= \left(\underset{\mathbf{X}\in\mathcal{S}_+}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{X}\mathbf{\Phi} - \overline{\mathbf{M}}\|_2^2\right)\mathbf{\Phi}, \end{aligned} \tag{4.25}$$

Note that $\mathcal{S}_2, \mathcal{S}_3$ and $\mathcal{S}_4$ can be handled through the definition of $\mathbf{\Phi}$. Let $\mathbf{M} = \mathbf{X}\mathbf{\Phi} \in \mathcal{B}_{\mathcal{S}_+}(\mathbf{\Phi})$ and $\overline{T} \in \{1, \ldots, K\}$ (K is the upper bound defined in (4.20)). Let $\mathbf{U} \in \mathbb{R}^{N\times\overline{T}}$ be a subpart of $\mathbf{X}$ with non-zero columns and $\mathbf{\Delta} \in \mathbb{C}^{\overline{T}\times L}$ the corresponding subpart of $\mathbf{\Phi}$ such that $\mathbf{M} = \mathbf{U}\mathbf{\Delta}$. Then we have

$$\mathcal{P}_{\mathcal{B}_{\mathcal{S}_+}(\mathbf{\Phi})}(\overline{\mathbf{M}}) = \Big(\underset{\mathbf{U}\in\mathbb{R}_+^{N\times\overline{T}}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{U}\mathbf{\Delta} - \overline{\mathbf{M}}\|_2^2\Big)\mathbf{\Delta}. \tag{4.26}$$

In (4.26), the dictionary $\mathbf{\Delta}$ is defined as

$$\mathbf{\Delta} = \underset{\overline{\mathbf{\Delta}}\in\mathcal{C}}{\operatorname{argmin}}\Big(\min_{\overline{\mathbf{U}}\in\mathbb{R}_+^{N\times\overline{T}}} \frac{1}{2}\|\overline{\mathbf{U}}\,\overline{\mathbf{\Delta}} - \overline{\mathbf{M}}\|_2^2\Big), \tag{4.27}$$

where $\mathcal{C}$ is the set given by

$$\mathcal{C} = \Big\{ \boldsymbol{\Delta} \in \mathbb{C}^{\overline{T} \times L} \,|\, (\exists \mathbf{X} \in \mathcal{S}_+)\ \mathbf{X} = \mathcal{Z}\big(\overline{\mathbf{U}}\big) \\ \text{with } \overline{\mathbf{U}} = \underset{\mathbf{U} \in \mathbb{R}_+^{N \times \overline{T}}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{U}\boldsymbol{\Delta} - \overline{\mathbf{M}}\|_2^2 \Big\}. \tag{4.28}$$

with $\mathcal{Z} \colon \mathbb{R}_+^{N \times \overline{T}} \to \mathbb{R}_+^{N \times D}$ defined such that $\mathcal{Z}(\mathbf{U})\boldsymbol{\Phi} = \mathbf{U}\boldsymbol{\Delta}$.

As mentioned earlier, $\boldsymbol{\Phi}$ is an over-complete dictionary which makes the exact projection practically impossible to compute. Recent advances in reconstruction methods have introduced neural networks to efficiently approximate the projection or proximal operators within model-based iterative algorithms [56–59]. A major challenge with such methods is obtaining sufficient accurate training data. In consequence, they can only accelerate the techniques where a prior computational solution to provide ground truth already exists. To overcome this difficulty, we propose a greedy approach to approximate the projection by finding a reduced dictionary $\widetilde{\boldsymbol{\Delta}} \in \mathbb{C}^{\overline{T} \times L}$ and its corresponding mixing matrix $\widetilde{\mathbf{U}} \in \mathbb{C}^{N \times \overline{T}}$, with $\overline{T} \leq K$, such that $\mathbf{U}\boldsymbol{\Delta} \approx \widetilde{\mathbf{U}}\widetilde{\boldsymbol{\Delta}}$. Then the projection in step 10 of Algorithm 6 can be approximated as

$$\mathcal{P}_{\mathcal{B}_{\mathcal{S}_+}(\boldsymbol{\Phi})}(\overline{\mathbf{M}}) \approx \left(\underset{\widetilde{\mathbf{U}} \in \mathbb{R}_+^{N \times \overline{T}}}{\operatorname{argmin}} \frac{1}{2} \left\| \widetilde{\mathbf{U}}\widetilde{\boldsymbol{\Delta}} - \overline{\mathbf{M}} \right\|_2^2 \right) \widetilde{\boldsymbol{\Delta}}. \tag{4.29}$$

As mentioned in [27], it is a common practice to allow the proton density to be complex-valued in order to absorb phase terms correcting for timing and coil sensitivity errors. We incorporate a vector $\boldsymbol{\lambda} \in \mathbb{C}^N$ to compensate for these errors. Let $\widetilde{\mathcal{B}}_{\mathcal{S}_+}(\boldsymbol{\Phi})$ be the set of magnetisation sequences of the form $\mathbf{M} = \operatorname{Diag}(\boldsymbol{\lambda})\mathbf{X}\boldsymbol{\Phi}$ such that $\mathbf{X} \in \mathcal{S}_+$ and $\boldsymbol{\lambda} \in \mathbb{C}^N$ satisfies $(\forall n \in \{1, \ldots, N\})\, |\boldsymbol{\lambda}_n| = 1$. The approximate projection with the phase compensation is given by:

$$\mathcal{P}_{\widetilde{\mathcal{B}}_{\mathcal{S}_+}(\boldsymbol{\Phi})}(\overline{\mathbf{M}}) \approx \operatorname{Diag}(\boldsymbol{\lambda})\widetilde{\mathbf{U}}\widetilde{\boldsymbol{\Delta}}, \tag{4.30}$$

where $(\boldsymbol{\lambda}, \widetilde{\mathbf{U}})$ are obtained by solving:

$$\begin{aligned} &\underset{\boldsymbol{\lambda}\in\mathbb{C}^N, \widetilde{\mathbf{U}}\in\mathbb{R}_+^{N\times\overline{T}}}{\text{minimise}} \frac{1}{2}\left\|\text{Diag}(\boldsymbol{\lambda})\widetilde{\mathbf{U}}\widetilde{\boldsymbol{\Delta}} - \overline{\mathbf{M}}\right\|_2^2 \\ &\text{subject to } (\forall n \in \{1,\ldots,N\})\, |\boldsymbol{\lambda}_n| = 1. \end{aligned} \tag{4.31}$$

It is worth mentioning that in (4.29) and (4.31), all the rows of $\widetilde{\mathbf{U}}$ can be computed independently in parallel. Note that the algorithms proposed in the previous chapter (i.e. ABLIP, ℓ_1-rABLIP and TV-rABLIP) can be explained using this framework. In particular, by constraining each row of $\overline{\mathbf{U}}$ to have at most one coefficient different than zero.

On the one hand, forward-backward based algorithms [18,60,61] can be used to solve problem (4.29) (in particular, in our simulations, we use the built-in MATLAB function of non-negative least-squares, that is an implementation of [62]). On the other hand, to solve problem (4.31), to jointly estimate $\boldsymbol{\Lambda}$ and $\mathbf{U}$, block coordinate approaches must be considered (e.g. Gauss-Seidel approaches [20, 63], alternating forward-backward methods [3, 45–47]). Note that in comparison with the traditional MRF methods which densely sample the manifold, our approach reduces the memory requirements, by using the dictionary $\widetilde{\boldsymbol{\Delta}}$ containing at most K elements, without the inaccuracies related to the manifold discretisation.

4.3.4 Greedy dictionary estimation

The GAP-MRF algorithm takes advantage of the dictionary coherence and the constraints imposed on $\mathbf{X}$ (described in Section 4.3.1) to approximate the projection onto $\mathcal{B}_{\mathcal{S}_+}(\boldsymbol{\Phi})$ in line 10 of Algorithm 6. As described in Section 4.3.3, this projection can be approximated at each iteration $i \in \mathbb{N}$, by solving (4.29), which necessitates to estimate the dictionary $\widetilde{\boldsymbol{\Delta}}^{(i)}$. We propose to estimate it using a greedy approach, leveraging both the knowledge of $\overline{\mathbf{M}}^{(i)}$ and the properties of the sets $\mathcal{S}_2$, $\mathcal{S}_3$ and $\mathcal{S}_4$ (note that the constraint $\mathcal{S}_1$ is handled directly in (4.29)). The proposed approach is described in details in this section.

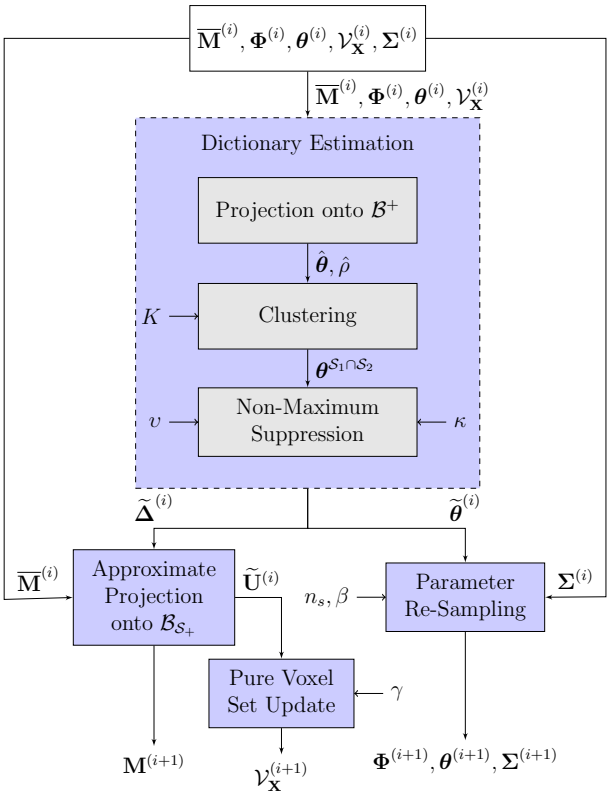


FIGURE 4.2: Greedy approximate projection diagram. The blue boxes represent the main steps in the approximate projection, the grey boxes represent the intermediate steps for the dictionary estimation and the arrows show the input and output variables.

The process to obtain $\widetilde{\mathbf{\Delta}}^{(i)}$ consists of three main steps leveraging the set of pure voxels. The first step consists in approximating the parameters of the pure voxels ($\mathcal{S}_4$ constraint) using the projection onto the set $\mathcal{B}^+$ defined in (3.20). The objective of the second step is to find K regions of interest ($\mathcal{S}_2$ constraint) of the manifold by exploiting its smoothness. Finally, in the third step, the parameters that are too close to each other are discarded ($\mathcal{S}_3$ constraint) by using a Non-Maximum Suppression based method [64]. This method acts on the number of voxels that corresponds to each parameter and keeps only the elements which have enough pure voxels to satisfy the $\mathcal{S}_4$ constraint. This process is summarised in the dictionary estimation step on Fig. 4.2. The remaining blue blocks in the diagram are used to update the variables in the greedy approximate projection. More precisely, we compute the mixing matrix $\widetilde{\mathbf{U}}^{(i)}$ and the magnetisation sequence $\mathbf{M}^{(i+1)}$ using equation (4.29) with the resulting dictionary $\widetilde{\mathbf{\Delta}}^{(i)}$. Then, we update the pure voxel set $\mathcal{V}_{\mathbf{X}}$ using the mixing matrix $\widetilde{\mathbf{U}}^{(i)}$.

Algorithm 7 Greedy Approximate Projection

Input: $\overline{\mathbf{M}}^{(i)}, \mathbf{\Phi}^{(i)}, \boldsymbol{\theta}^{(i)}, \mathcal{V}_{\mathbf{X}}^{(i)}, \mathbf{\Sigma}^{(i)}, K, \mathbf{\Gamma}, \kappa, \upsilon, \gamma, \beta, \xi, n_s$

Dictionary Estimation:

Projection onto $\mathcal{B}^+$

for $n = 1, 2, ..., N$ **do**

$\hat{d}_n = \underset{d}{\text{argmax}}\ \text{real}(\overline{\mathbf{M}}_{n,:}^{(i)} \mathbf{\Phi}_{d,:}^{\dagger(i)}) / \|\mathbf{\Phi}_{d,:}^{(i)}\|_2$

$\hat{\rho}_n = \max(\text{real}(\overline{\mathbf{M}}_{n,:}^{(i)} \mathbf{\Phi}_{\hat{d}_n,:}^{\dagger(i)}) / \|\mathbf{\Phi}_{\hat{d}_n,:}^{(i)}\|_2^2, 0)$

$\hat{\boldsymbol{\theta}}_{n,:} = \boldsymbol{\theta}_{\hat{d}_n,:}^{(i)}$

Clustering

$\mathcal{I} = \{n \in \mathcal{V}_{\mathbf{X}}^{(i)} \mid \hat{\rho}_n > \xi\}$

$[\boldsymbol{\theta}^{\mathcal{S}_1 \cap \mathcal{S}_2}, \mathbf{c}] = \text{k-means}(\hat{\boldsymbol{\theta}}_{\mathcal{I},:}, K)$

Non-Maximum Suppression

$\widetilde{\boldsymbol{\theta}}^{(i)} = \text{NonMaximumSuppression}(\boldsymbol{\theta}^{\mathcal{S}_1 \cap \mathcal{S}_2}, \mathbf{c}, \upsilon, \kappa)$

$\widetilde{\mathbf{\Delta}}^{(i)} = B(\widetilde{\boldsymbol{\theta}}^{(i)}, \mathbf{\Gamma})$

Approximate Projection onto $\mathcal{B}_{\mathcal{S}_+}$

$\widetilde{\mathbf{U}}^{(i)} = \underset{\overline{\mathbf{U}} \in \mathbb{R}_+^{N \times T}}{\text{argmin}} \frac{1}{2} \|\overline{\mathbf{U}} \widetilde{\mathbf{\Delta}}^{(i)} - \overline{\mathbf{M}}^{(i)}\|_2^2$

$\mathbf{M}^{(i+1)} = \widetilde{\mathbf{U}}^{(i)} \widetilde{\mathbf{\Delta}}^{(i)}$

Pure Voxel Set Update

$\mathcal{G}_{\mathbf{X}}^{(i)} = \{n \in \{1, \ldots, N\} \mid \|\widetilde{\mathbf{U}}_{n,:}^{(i)}\|_1 > \xi\}$

$\mathcal{V}_{\mathbf{X}}^{(i+1)} = \{n \in \mathcal{G}_{\mathbf{X}}^{(i)} \mid \max(\widetilde{\mathbf{U}}_{n,:}^{(i)}) \geq \gamma \|\widetilde{\mathbf{U}}_{n,:}^{(i)}\|_1\}$

Parameter Re-sampling

$\boldsymbol{\theta}^{(i+1)} = \text{ParameterReSampling}(\widetilde{\boldsymbol{\theta}}^{(i)}, \mathbf{\Sigma}^{(i)}, n_s)$

$\mathbf{\Phi}^{(i+1)} = B(\boldsymbol{\theta}^{(i+1)}, \mathbf{\Gamma})$

$\mathbf{\Sigma}^{(i+1)} = \mathbf{\Sigma}^{(i)} \beta$

Output: $\boldsymbol{\theta}^{(i+1)}, \mathbf{\Phi}^{(i+1)}, \mathbf{\Sigma}^{(i+1)}, \mathcal{V}_{\mathbf{X}}^{(i+1)}$ and $\mathbf{M}^{(i+1)}$

Finally, the dictionary $\mathbf{\Phi}$ is refined by randomly sampling around the parameters $\widetilde{\boldsymbol{\theta}}^{(i)}$. The complete method is described in Algorithm 7 and explained in the following paragraphs.

4.3.4.1 Projection onto $\mathcal{B}^+$

At iteration $i \in \mathbb{N}$, we have:

$$\mathcal{P}_{\mathcal{B}_{\mathcal{S}_+(\mathbf{\Phi})}}\left(\overline{\mathbf{M}}_{\mathcal{V}_{\mathbf{X}}^{(i)},:}\right) = \mathcal{P}_{\mathcal{B}^+(\mathbf{\Phi})}\left(\overline{\mathbf{M}}_{\mathcal{V}_{\mathbf{X}}^{(i)},:}\right), \tag{4.32}$$

where $\mathcal{B}^+$ is the set defined in equation (3.20), and $\overline{\mathbf{M}}_{\mathcal{V}_{\mathbf{X}}^{(i)},:} = (\overline{\mathbf{M}}_{n,:})_{n\in\mathcal{V}_{\mathbf{X}}^{(i)}}$, $\mathcal{V}_{\mathbf{X}}^{(i)}$ corresponding to an estimate of the pure voxel positions in $\mathbf{X}^{(i)}$ at iteration i (the true set $\mathcal{V}_{\mathbf{X}}$ corresponding to the pure voxels of the original $\mathbf{X}$ being unknown). At the first iteration, we choose $\mathcal{V}_{\mathbf{X}}^{(0)} = \{1, \ldots, N\}$, and it is updated during the greedy process (see Algorithm 7, step 19). Note that that the region containing pure voxel does not need to be known *a priori* (it is automatically estimated), and does not need to be large to be detected (depending on the noise).

From (4.32), we can estimate the parameters $\hat{\boldsymbol{\theta}}$ and the proton density $\hat{\rho}$ of the voxels in $\mathcal{V}_{\mathbf{X}}^{(i)}$ using the projection onto $\mathcal{B}^+$ with a dictionary $\mathbf{\Phi}^{(i)}$ (see steps 4-7 of Algorithm 7). $\mathbf{\Phi}^{(i)}$ is an adaptive dictionary that is refined at each iteration to reduce the computational cost, the simulations suggest that the accuracy of the reconstructions is preserved. Since there are at least κ pure voxels for each active element in $\mathbf{\Phi}$ and the value of the proton density is at least ξ, we expect that the voxel parameters in $\mathcal{V}_{\mathbf{X}}^{(i)}$ with $\hat{\rho} > \xi$ will form clusters around the true values of the dictionary elements, an example can be seen in Fig. 4.3 (Left).

4.3.4.2 Clustering

In order to find K centres approximating the parameters of interest, we propose to use the k-means algorithm [65]. The objective of K-means is to find K centres that minimise the squared distance from all points to its closest centre. The centres obtained by solving the k-means problem $\boldsymbol{\theta}^{\mathcal{S}_1\cap\mathcal{S}_2} \in \mathbb{R}^{K\times P}$ can be used to compute a dictionary $\mathbf{\Delta}^{\mathcal{S}_1\cap\mathcal{S}_2} \in \mathbb{C}^{K\times L}$. By solving equation (4.29) with $\mathbf{\Delta}^{\mathcal{S}_1\cap\mathcal{S}_2}$, we would obtain a $\mathbf{U}^{\mathcal{S}_1\cap\mathcal{S}_2} \in \mathbb{R}^{N\times K}$ such that $\mathcal{Z}\left(\mathbf{U}^{\mathcal{S}_1\cap\mathcal{S}_2}\right) \in \mathcal{S}_1 \cap \mathcal{S}_2$.

4.3.4.3 Non-maximum suppression

The k-means algorithm also provides a label to each voxel corresponding to the matched centre. We define $\mathbf{c} \in \mathbb{R}^{K\times 1}$ to be the vector containing the number of voxels associated with each centre. Inspired by the Non-Maximum Suppression method in [64], we use the number of pure voxels assigned to each centre to remove the neighbours defined in equation (4.21). We first take the parameters of the highest value of $\mathbf{c}$, and we add all the $\mathbf{c}$ values of the neighbours to the maximum value of $\mathbf{c}$ if it is greater than κ we keep the parameters, if not we discard them and set the corresponding values of $\mathbf{c}$ to 0 (see Fig. 4.3 (centre)). We repeat the process until all values of $\mathbf{c}$ are 0. Finally, we use the resulting parameters $\widetilde{\boldsymbol{\theta}}^{(i)} \in \mathbb{R}^{\overline{T}\times P}$ to construct $\widetilde{\boldsymbol{\Delta}}^{(i)} \in \mathbb{C}^{\overline{T}\times L}$.

4.3.4.4 Inexact projection onto $\mathcal{B}_{\mathcal{S}_+}$

Once the dictionary $\widetilde{\boldsymbol{\Delta}}^{(i)}$ is approximated, computing the three steps described above, the magnetisation sequence $\mathbf{M}^{(i+1)}$ can be updated. To this aim, we use equation (4.29), where the minimisation problem is solved using MATLAB built-in function for non-negative least-squares problems [62].

4.3.4.5 Pure voxel set update

In order to avoid noisy voxels, we re-define the set $\mathcal{G}_{\mathbf{X}}$, introduced in Section 4.3.1.2, for $\widetilde{\mathbf{U}}^{(i)}$. Note that $\mathcal{Z}\big(\widetilde{\mathbf{U}}^{(i)}\big)$ is a matrix of the size of $\mathbf{X}$ filling the missing values of $\widetilde{\mathbf{U}}^{(i)}$ with zeros, and thus we can re-define the set $\mathcal{G}_{\mathbf{X}}^{(i)}$ in terms of $\widetilde{\mathbf{U}}^{(i)}$ as:

$$\mathcal{G}_{\mathbf{X}}^{(i)} = \{n \in \{1,\dots,N\} \,|\, \|\widetilde{\mathbf{U}}_{n,:}^{(i)}\|_1 > \xi\}. \tag{4.33}$$

Then, we update the pure voxel set as:

$$\mathcal{V}_{\mathbf{X}}^{(i+1)} = \{n \in \mathcal{G}_{\mathbf{X}}^{(i)} \,|\, \max\left(\widetilde{\mathbf{U}}_{n,:}^{(i)}\right) \geq \gamma \|\widetilde{\mathbf{U}}_{n,:}^{(i)}\|_1\}, \tag{4.34}$$

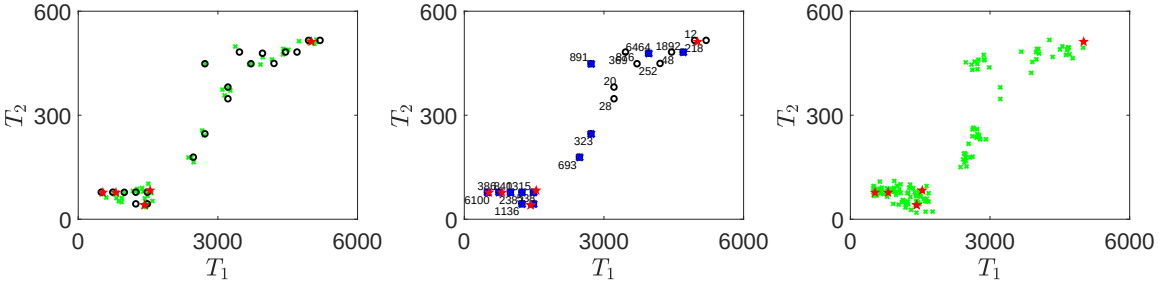


FIGURE 4.3: Examples of the clustering, non-maximum suppression and parameter re-sampling. For all examples, the red stars represent the true phantom parameters. (Left) Clustering. The parameters of the voxels in $\mathcal{V}_{\mathbf{X}}^{(i)}$ which its corresponding proton density is greater than ξ (green crosses) are the input of the k-means algorithm and the output are the centres (black circles). (centre) Non-maximum suppression. The centres obtained by the k-means (black circles) and the filtered centres are the output of the Non-Maximum Suppression (blue crosses). (Right) Parameter re-sampling. The parameters of the dictionary $\mathbf{\Phi}^{(i+1)}$ are obtained by randomly sampling around the parameters obtained by the Non-Maximum Suppression (green crosses).

where $0 < \gamma < 1$ is a relaxation factor used to compensate both for the noise and for the fact that the true dictionary elements are not guaranteed to be present. Note that the parameter γ is defined as a proportion of the total proton density in the voxels, and it is used as a threshold to determine if a voxel is pure or not.

4.3.4.6 Parameter re-sampling

We update $\mathbf{\Phi}^{(i)}$ to refine the manifold elements of interest. For this process, we produce n_s random samples around the elements in $\widetilde{\boldsymbol{\theta}}^{(i)}$ using a Gaussian distribution with a diagonal covariance matrix $\mathbf{\Sigma}^{(i)}$ (see Fig. 4.3 (Right)). The values of the covariance matrix $\mathbf{\Sigma}^{(i)}$ are reduced by a factor $0 < \beta < 1$ at each iteration. When the values of $\mathbf{\Sigma}^{(i)}$ are sufficiently small, the dictionary $\widetilde{\mathbf{\Delta}}$ will not change anymore and after a fixed number of iterations the sequences generated by Algorithm 6 will stabilise. Since the samples are randomly Gaussian distributed, the parameter values are not limited to a given resolution.

4.4 Choice of the parameters and initialisation

Since $\mathcal{S}_+$ is a non-convex set, the choice of the initialisation is important. If the initial magnetisation sequence or the dictionary are not close to the desired values, the greedy approximate projection can fail. In this section, we will describe the initialisation for our algorithm.

4.4.1 Choice of the parameters

The choice of ξ, setting the minimum proton density, is related to the background noise, the ideal ξ is a value between the background noise and the signal in the volume of interest. If ξ is too small, empty voxels will affect the clustering process. If it is too big, the tissue voxels will not be considered in the clustering process.

As mentioned before, the dictionary $\mathbf{\Phi}^{(i)}$ is updated through the iterations to reduce the complexity of the algorithm. We fix $\mathbf{\Phi}^{(0)}$ to all possible combinations of 20 values of T_1 and 20 values of T_2, equally spaced in $\mathcal{M}$.

Concerning the number of random samples n_s, on the one hand, if we choose it too big, we increase the complexity of our pure voxel projection. On the other hand, if we set n_s too small, more iterations will be needed to find the elements of interest. In all our simulations (simulated and *in vivo* data) we fix $n_s = 10$.

For the diagonal elements of the covariance matrix (i.e. $\mathbf{\Sigma}^{(0)}_{1,1}$ and $\mathbf{\Sigma}^{(0)}_{2,2}$) associated to the resampling of the dictionary, if they are chosen too big, the parameter sampling will be far from the parameters of interest, increasing the number of iterations required to find them. If they are too small, the algorithm may not find the parameter of interest. $\mathbf{\Sigma}^{(0)}$ should be chosen based on the parameter separation of $\mathbf{\Phi}^{(0)}$. In all the reconstructions we fix $\mathbf{\Sigma}^{(0)}_{1,1} = 40$ and $\mathbf{\Sigma}^{(0)}_{2,2} = 10$.

Similarly, for the decreasing parameter β of the covariance matrix (see step 23 in Algorithm 7), if it is chosen too big, the algorithm will need more iterations to find

the correct elements while if it is too small the algorithm may not explore the true parameters. We fix $\beta = 0.9$ in the considered scenarios.

The choice of the pure voxel tolerance γ is related to the noise and the accuracy of the dictionary during the iterations of the algorithm. If it is too big, the elements of interest could be eliminated through the iterations since pure voxels may be considered as PV voxels, if it is too small, the PV voxels may be considered as pure affecting the clustering process. We found in our simulations that $\gamma = 0.85$ is a suitable choice.

The choice of the different parameters K, υ and κ has been investigated during preliminary work. In particular, we observed a significant increase in the residual $\|\mathbf{Y} - h(\mathbf{M})\|_2$ when K is not sufficiently large. For υ and κ, we see a significant increase in the residual when they are chosen too large (i.e. merging proton density maps of the true tissues), and an increase of noisy proton density maps when they are chosen too low.

We propose to automatically choose K, υ and κ by analysing the residual. Precisely, we choose a tolerance value on the residual, denoted by $\tau > 0$. This value, indicates the minimum contribution of an element of the dictionary in the residual. If τ is chosen too big, our solution will contain noisy elements. While if it is chosen too small our elements of interest will be removed from the reconstruction.

4.4.2 Initialisation

The global GAP-MRF method, including the initialisation process, is described in Algorithm 8. It describes the process to choose the parameters K, υ and κ. Firstly, the estimation of K is described in steps 2-10. Fixing all the other parameters, K is estimated by running multiple times the GAP-MRF iterations given in Algorithm 1. We assume that we have a suitable estimate of K when the stopping criteria given in step 10 of Algorithm 8 is reached. The same process is adopted for the estimation of υ described (steps 11-19) and κ (steps 20-28). For these two estimates, we allow for a small tolerance ($\tau_\upsilon > 0$ and $\tau_\kappa > 0$, respectively), for robustness purposes. Note that

Algorithm 8 GAP-MRF global method

Input: $\mathbf{Y}$, $\mathbf{\Gamma}$, $\mathbf{\Phi}$, $\boldsymbol{\theta}$, ξ, τ, $\mathbf{\Sigma}_{1,1}^{(0)} = 40$, $\mathbf{\Sigma}_{2,2}^{(0)} = 10$, $\zeta = 0.99$, $\mathcal{V}_{\mathbf{X}}^{(0)} = \{1, \ldots, N\}$, $\beta = 0.9$, $n_s = 10$, $\mathbf{M}^{(0)} = \mathbf{0}$, $(\tau_K, \tau_\upsilon, \tau_\kappa) = (10, 0.02, 10)$

Estimation of K:

  Input: $(\gamma, K, \upsilon, \kappa) = (0, 0, 0, 0)$, $\widetilde{\mathbf{\Delta}}^{(0)} = \{\}$, $\widetilde{\boldsymbol{\theta}}^{(0)} = \{\}$, $j = 0$.

  Do

    $\mathbf{\Phi}^{(0)} = \left[\mathbf{\Phi}, \widetilde{\mathbf{\Delta}}^{(j)}\right]$, $\boldsymbol{\theta}^{(0)} = \left[(\boldsymbol{\theta})^T, (\widetilde{\boldsymbol{\theta}}^{(j)})^T\right]^T$

    $K = K + \tau_K$

    $\left[\mathbf{M}^{(j+1)}, \widetilde{\mathbf{\Delta}}^{(j+1)}, \widetilde{\boldsymbol{\theta}}^{(j+1)}\right] =$ Algorithm6$(\mathbf{Y}, \zeta, \mathbf{M}^{(j)})$

    j=j+1

  while $\|\mathbf{Y} - h(\mathbf{M}^{(j-1)})\|_2 - \|\mathbf{Y} - h(\mathbf{M}^{(j)})\|_2 > \tau$.

  Output: $K^\star = K$, $j = j - 1$

Estimation of υ:

  Input: $(\gamma, K, \upsilon, \kappa) = (0.85, K^\star, 0, 0)$

  Do

    $\mathbf{\Phi}^{(0)} = \widetilde{\mathbf{\Delta}}^{(j)}$, $\boldsymbol{\theta}^{(0)} = \widetilde{\boldsymbol{\theta}}^{(j)}$

    $\upsilon = \upsilon + \tau_\upsilon$

    $\left[\mathbf{M}^{(j+1)}, \widetilde{\mathbf{\Delta}}^{(j+1)}, \widetilde{\boldsymbol{\theta}}^{(j+1)}\right] =$ Algorithm6 $\left(\mathbf{Y}, \zeta, \mathbf{M}^{(j)}\right)$

    j=j+1;

  while $\|\mathbf{Y} - h(\mathbf{M}^{(j)})\|_2 - \|\mathbf{Y} - h(\mathbf{M}^{(j-1)})\|_2 > \tau$.

  Output: $\upsilon^\star = \upsilon - 2\tau_\upsilon$, $j = j - 1$

Estimation of κ:

  Input: $(\gamma, K, \upsilon, \kappa) = (0.85, K^\star, \upsilon^\star, 0)$

  Do

    $\mathbf{\Phi}^{(0)} = \widetilde{\mathbf{\Delta}}^{(j)}$, $\boldsymbol{\theta}^{(0)} = \widetilde{\boldsymbol{\theta}}^{(j)}$

    $\kappa = \kappa + \tau_\kappa$

    $\left[\mathbf{M}^{(j+1)}, \widetilde{\mathbf{\Delta}}^{(j+1)}, \widetilde{\boldsymbol{\theta}}^{(j+1)}\right] =$ Algorithm6 $\left(\mathbf{Y}, \zeta, \mathbf{M}^{(j)}\right)$

    j=j+1;

  while $\|\mathbf{Y} - h(\mathbf{M}^{(j)})\|_2 - \|\mathbf{Y} - h(\mathbf{M}^{(j-1)})\|_2 > \tau$.

  Output: $\kappa^\star = \kappa - 2\tau_\kappa$

GAP-MRF Global Iterations:

  Input: $(\gamma, K, \upsilon, \kappa) = (0.85, K^\star, \upsilon^\star, \kappa^\star)$, $\mathbf{\Phi}^{(0)} = \widetilde{\mathbf{\Delta}}^{(j-1)}$, $\boldsymbol{\theta}^{(0)} = \widetilde{\boldsymbol{\theta}}^{(j-1)}$

  $\left[\overline{\mathbf{M}}, \overline{\mathbf{\Delta}}, \overline{\boldsymbol{\theta}}\right] =$ Algorithm6 $\left(\mathbf{Y}, \zeta, \mathbf{M}^{(j-1)}\right)$

  $\overline{\mathbf{U}} = \underset{\mathbf{U} \in \mathbb{R}_+^{N \times \overline{T}}}{\operatorname{argmin}} \dfrac{1}{2} \|\mathbf{U}\overline{\mathbf{\Delta}} - \overline{\mathbf{M}}\|_2^2$

  Output: $\overline{\mathbf{M}}$, $\overline{\mathbf{\Delta}}$, $\overline{\boldsymbol{\theta}}$ and $\overline{\mathbf{U}}$

each new run of Algorithm 6 uses the previous estimated of $\mathbf{M}$, $\mathbf{\Delta}$ and $\boldsymbol{\theta}$, in order to accelerate the global method.

4.5 Results

We create a simulated phantom according to [33], with five tissues: adipose, WM, GM, muscle and cerebrospinal fluid (CSF). More precisely, to introduce the PVE, we use blocks of 2×2 voxels to form a lower resolution phantom containing PV voxels. The resulting volume is resized to 256×256 voxels, and we generate the magnetisation sequence from this volume using the Bloch equations. All the reconstructions are performed with the same resolution. In the first column of Fig. 4.4 proton density maps and the voxel distribution of the simulated phantom are shown. Using this representation we can see the structure of the tissues of interest. Traditionally in qMRI, individual parameter maps are evaluated since only a tissue per voxel is considered but in a PV scenario this is not meaningful since several parameter maps would be needed and visually do not show the tissue structures. We also compute the dominant tissue (highest proton density in the voxel) parameter maps for a traditional evaluation. The phantom dominant tissue parameter maps can be seen in the first column of Fig. 4.6. Note that for the construction of the phantom, we only consider in-plane PV, while in reality through-plane PV and in-plane PV occurs. Both kind of PV are modelled the same way and should not make any difference in the reconstructions.

In order to evaluate the algorithms, we use the Signal-to-Noise Ratio (SNR in dB) defined as

$$\mathrm{SNR}(\mathbf{U}_{:,t}, \widetilde{\mathbf{U}}_{:,t}) = 10 \log \left(\frac{\sum_{n=1}^{N} (\mathbf{U}_{n,t})^2}{\sum_{n=1}^{N} (\mathbf{U}_{n,t} - \widetilde{\mathbf{U}}_{n,t})^2} \right), \tag{4.35}$$

where $t \in \{1, \ldots, T\}$ is the index of the evaluated tissue, $\mathbf{U}$ is the mixing matrix ground truth and $\widetilde{\mathbf{U}}$ is the estimation. Similarly for the magnetisation sequence SNR, we sum for all values in the matrix. To construct the matrix $\widetilde{\mathbf{U}}$, a tolerance of 15% from the ground truth parameter values is used (i.e. for $T_1 = 530$ and $T_2 = 77$ milliseconds (ms) all the dictionary elements that fall for T_1 in the range of $[450.5 - 609.5]$ms and simultaneously for T_2 in the range of $[65.45 - 88.55]$ms are considered). In order to evaluate if the tissues are correctly identified, we define the success rate (SR) index as the proportion of voxels where the number of elements are correctly identified and its corresponding parameters fall within the 15% of the true parameters. The

same definition of SR is used for both pure and PV voxels (considering only the corresponding phantom voxels). Due to noise, there could be small values in $\widetilde{\mathbf{U}}$ that could significantly affect the SR. In consequence, we choose not to consider values that are smaller than 30, given that the range of the proton densities is from 80 to 400.

4.5.1 Experiment 1 - Impact of the iSNR

In this experiment, we investigate the behaviour of both the IR-MRF and the GAP-MRF algorithms while changing the input noise. We fix the magnetisation sequence length $L = 1000$. The dictionary for BLIP is defined as in [27] with $D = 16170$. The results correspond to an average (with standard deviation) over 10 runs of each choice of iSNR.

The results of the proton density maps are shown in Fig. 4.7 (Left). GAP-MRF significantly outperform BLIP when the iSNR is greater than 30dB. We can notice that GAP-MRF estimates correctly the number of true atoms when the iSNR is 30dB or greater. The reconstruction of adipose tissue is more affected by the noise since there are significantly less pure voxels of this tissue. GAP-MRF magnetisation sequence reconstruction is significantly more accurate than BLIP reconstruction, because BLIP does not consider the PVE and also because of the dictionary inaccuracy. GAP-MRF magnetisation sequence SNR has a linear behaviour with respect to the iSNR. In Fig. 4.7 (Center), the SR with respect to the iSNR can be seen. We can observe that the SR is significantly affected by the iSNR.

The results for the dominant tissue parameter maps SNR can be seen in Fig. 4.7 (Right). GAP-MRF outperforms BLIP reconstructing the dominant tissue parameter maps. It is important to mention than GAP-MRF is more affected by noise because the linear combination of dictionary elements overfits the noise.

We show an example of the proton density maps for each tissue in Fig. 4.4 when the iSNR is 30dB. By visual inspection, we can observe that the GAP-MRF method outperforms the BLIP method for PV reconstructions for moderate noise scenarios.

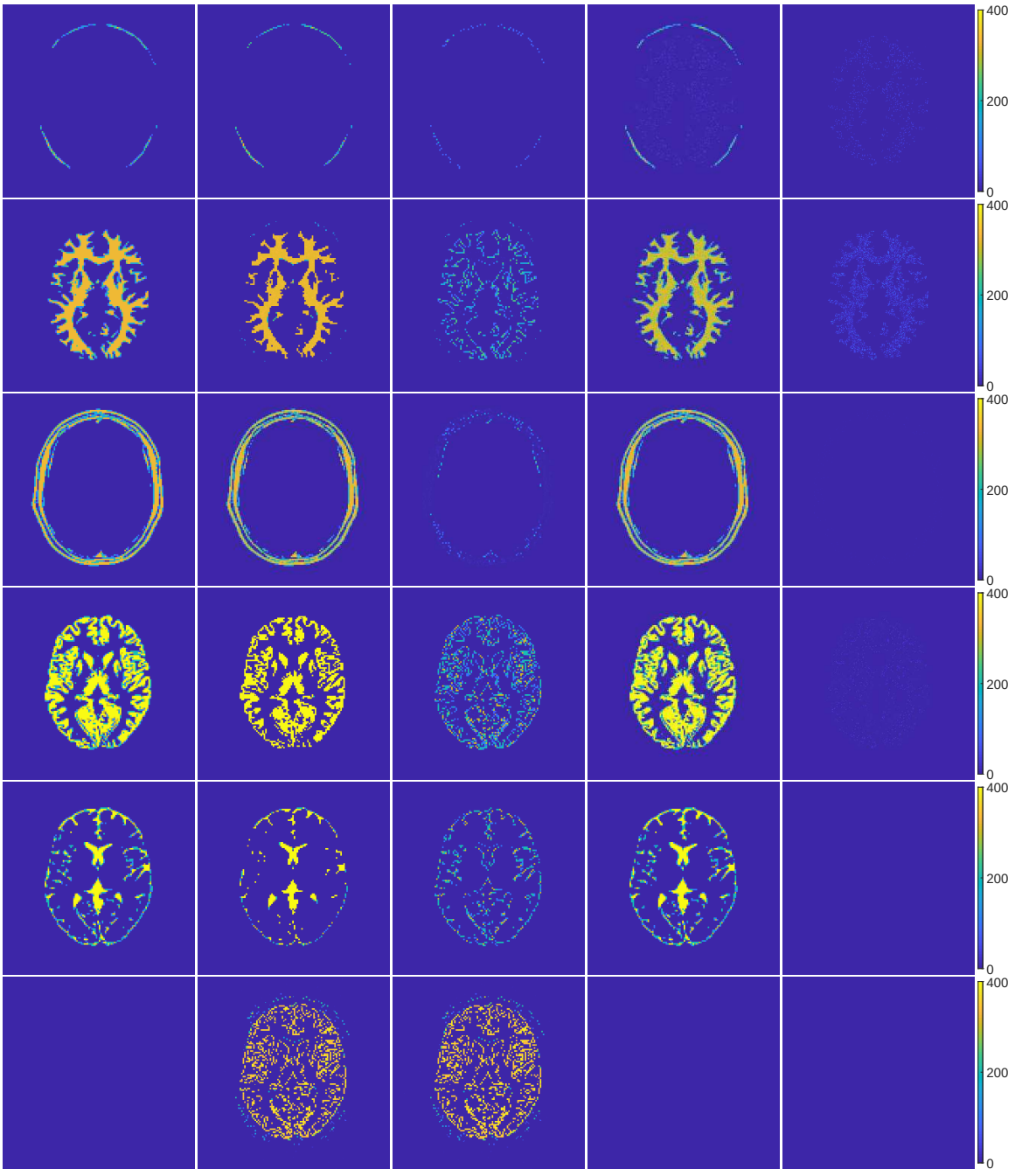


FIGURE 4.4: Experiment 1 - Example of the proton density maps with $L = 1000$ and an iSNR of 30dB. From first to the last column: Ground truth images, BLIP reconstructions, absolute difference between BLIP and the Ground Truth, GAP-MRF reconstructions and absolute difference between GAP-MRF and the Ground Truth. From first to the fifth row: Adipose, WM, muscle, GM and CSF. Sixth row: Proton density sum of all other matched elements that are not in the 15% range of the ground truth elements. The corresponding T_1 and T_2 values are given in Table 4.1.

The values of BLIP in Table 4.1 are given in a range because multiple parameters were assigned to the corresponding ground truth tissue. On the contrary, GAP-MRF has

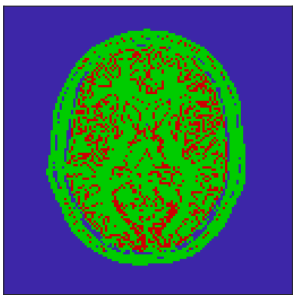


FIGURE 4.5: Voxel distribution map of the simulated phantom showing the pure voxels (green) and the PV voxels (red).

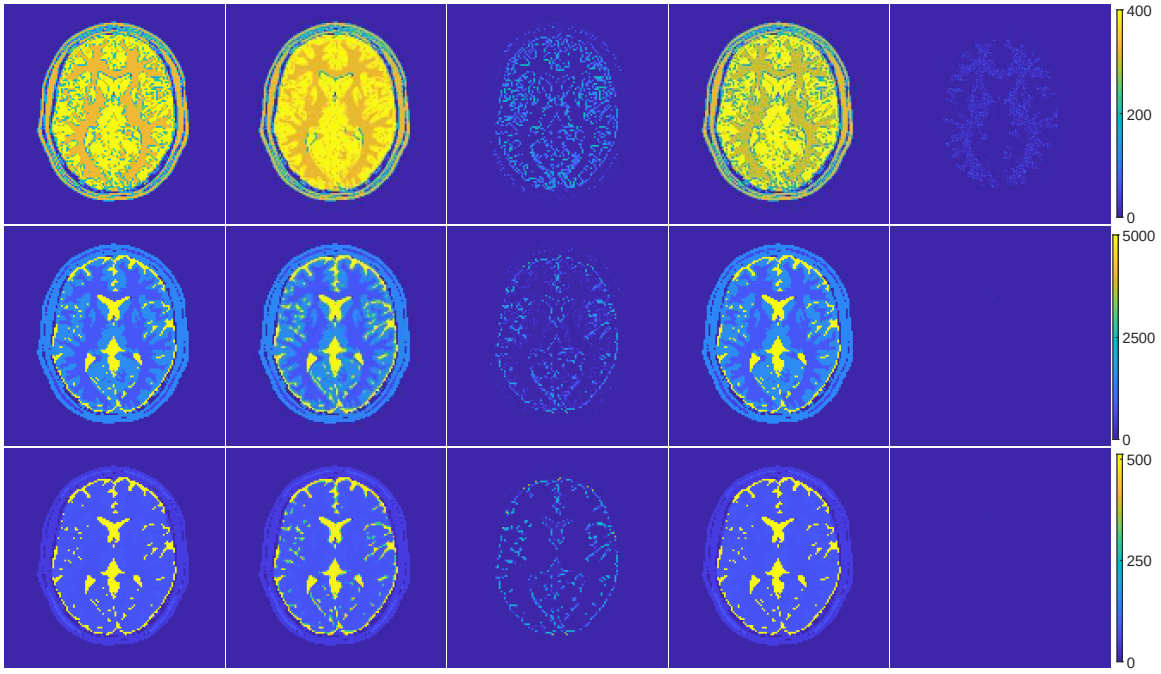


FIGURE 4.6: Experiment 1 - Example of the dominant tissue parameters with $L = 1000$ and an iSNR of 30dB. From first to the last column: Ground truth images, BLIP reconstructions, absolute difference between BLIP and the Ground Truth, GAP-MRF reconstructions and absolute difference between GAP-MRF and the Ground Truth. From first to the last row: Proton density, T_1 and T_2 parameter maps.

a single value because only one value was assigned to the corresponding ground truth tissue. In this example, for BLIP and GAP-MRF respectively, the SNR values are as follows: 9.70dB and 11.94dB for Adipose, 9.14dB and 19.52dB for WM, 17.66dB and 39.29dB for Muscle, 8.31dB and 31.87dB for GM, 5.72dB and 52.60dB for CSF and for the magnetisation sequence 23.84dB and 48.18dB. The SR: 0.9944 and 0.9745 for pure voxels, and for PV voxels 0 and 0.9465. The GAP-MRF correctly estimates the manifold regions of interest. BLIP has a residual map formed by all the elements

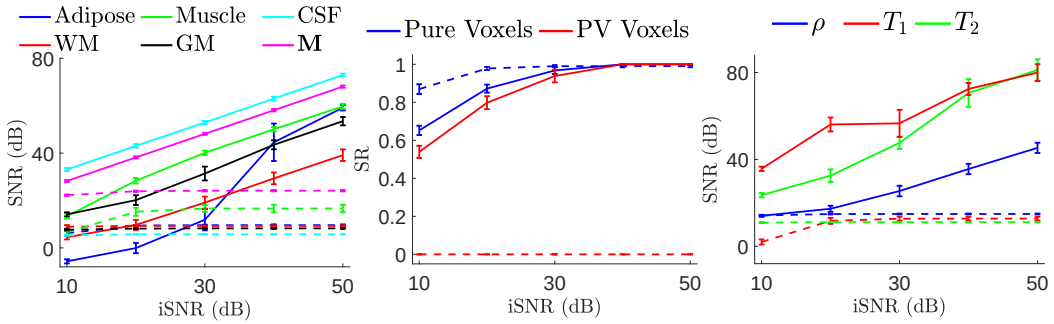


FIGURE 4.7: Experiment 1 - Simulation results obtained with BLIP (dashed lines) and GAP-MRF (solid lines). Left: Tissue proton density maps (Adipose, WM, Muscle, GM, CSF) and magnetisation sequence ($\mathbf{M}$) evaluation. Center: SR evaluation for the pure and PV voxels. Right: Dominant tissue parameter maps (ρ, T_1, T_2) evaluation.

TABLE 4.1: Parameter values of example in Fig. 4.4 corresponding to Experiment 1 with an iSNR of 30dB. The relaxation times are in ms.

	Ground Truth		BLIP		GAP-MRF	
	T_1	T_2	T_1	T_2	T_1	T_2
Adipose	530	77	[460-590]	[74-84]	531.1	77.0
White Matter	811	77	[690-930]	[66-80]	811.1	77.0
Muscle	1425	41	[1220-1630]	[36-46]	1424.0	41.0
Gray Matter	1545	83	[1320-1610]	[74-86]	1544.3	83.1
CSF	5012	512	[4400-5000]	500	5013.1	512.1

that are not sufficiently close to the true elements (see last row of Fig. 4.4). Note that the residual map is quite similar to the distribution of the PV voxels shown in Fig. 4.5, this shows that the parameter mismatch is due to the PVE. In the GAP-MRF reconstructions, the WM and adipose tissue are slightly mixed due to the noise since their parameters are close one to each other. By choosing a better $\mathbf{\Gamma}$ we can make the atoms of the dictionary more distant in the ℓ_2-norm sense, this would provide noise robustness to the reconstructions. The dominant tissue parameter maps are shown in Fig. 4.6. The T_1 and T_2 maps reconstructed by BLIP show a smooth transition from one tissue to another due to the partial volume. On the contrary, GAP-MRF reconstructions show abrupt transitions in the T_1 and T_2 maps delimiting the tissues. This is expected since each tissue is modelled with a unique set of parameters. We can observe that the dominant tissue proton density reconstruction of GAP-MRF is significantly affected by the noise. Nevertheless, thanks to the constraint $\mathcal{S}_+$ handled by the proposed method, the T_1 and T_2 parameter maps are accurate.

4.5.2 Experiment 2 - Impact of the magnetisation sequence length

In this subsection, we compare the proposed GAP-MRF algorithm with the BLIP algorithm, for different number of excitation instances L. The iSNR is set to 50dB. The dictionary for BLIP is defined as in [27] with $D = 16170$. The results correspond to an average (with standard deviation) over 10 runs of each choice of L.

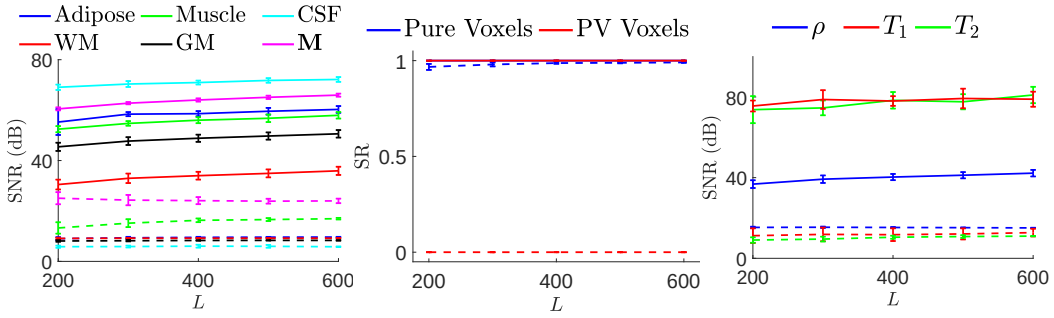


FIGURE 4.8: Experiment 2 - Simulation results obtained with BLIP (dashed lines) and GAP-MRF (solid lines). Left: Tissue proton density maps (Adipose, WM, Muscle, GM, CSF) and magnetisation sequence ($\mathbf{M}$) evaluation. Center: SR evaluation for the pure and PV voxels. Right: Dominant tissue parameter maps (ρ, T_1, T_2) evaluation.

Fig. 4.8 (Left) shows the evaluation of the proton density maps for each tissue (Adipose, WM, Muscle, GM, and CSF) and the magnetisation sequence. Note that GAP-MRF results are taken directly from the matrix $\widetilde{\mathbf{U}}$ without using any post-processing. We can observe that GAP-MRF outperforms BLIP in reconstructing $\mathbf{U}$. This can be explained by the fact that BLIP is restricted to the input dictionary, while our method estimates the dictionary. In addition, we can observe that the SNR values of magnetisation sequence reconstructed with BLIP slightly decreases while L increases (while it is not the case for the proton density maps). This is expected since the linear combination of short fingerprints are less distinctive, hence it is easier to approximate it with fingerprints of other elements (allowing BLIP to fit better PV voxels with other elements). This behaviour is not observed with the proposed GAP-MRF method for which accurate proton density map estimates result in accurate magnetisation sequence reconstructions.

Fig. 4.8 (Center) gives the SR for both pure and PV voxels. Since BLIP can only reconstruct one element per voxel, its SR for PV is always equal to 0. In a low noise scenario, GAP-MRF can identify the correct voxel elements even for short sequences as can be seen in Fig. 4.8 (Center), where the SR of GAP-MRF for pure and PV voxels is 1. An important remark is that due to PV, the dictionary sampling and the number of excitation instances, the BLIP algorithm can mis-reconstruct pure voxels even in a low noise scenario.

For the dominant tissue parameter maps in the low noise scenario, GAP-MRF outperforms BLIP as shown in Fig. 4.8 (Right). The proton density map of BLIP is affected by the PV since it is not able to distinguish between the voxel tissues. The T_1 and T_2 maps are affected by the PV voxels and the dictionary inaccuracies.

BLIP reconstructions show a variation on T_1 and T_2 for the same tissue while GAP-MRF reconstructions are accurate. The GAP-MRF has the additional advantage that it simultaneously estimates the manifold regions of interest, resulting in better reconstructions.

4.5.3 Real data results

In this section, we show the reconstructions on the EUROSPIN phantom and on two *in vivo* datasets. The first and second datasets were acquired using spiral sampling scheme and the third dataset was acquired using EPI sampling scheme [66]. The parameters were chosen as discussed in Section 4.4. The obtained proton density maps were normalised as $\widetilde{\mathbf{U}}/\max(\widetilde{\mathbf{U}})$ and only the proton densities greater than the 10% of $\max(\widetilde{\mathbf{U}})$ are shown in the figures. The normalised proton density is in arbitrary units (a.u.) and the relaxation times are in ms. Note that for the spiral datasets, a single spiral interleaf is acquired for each excitation instance. For the next excitation instance, the interleaf is rotated a fixed angle given by the total number of interleaves (e.g. 377 interleaves corresponds to $360/377 \approx .9549°$).

4.5.3.1 EUROSPIN phantom dataset with spiral sampling

In this subsection, we show the results obtained with the proposed approach and BLIP method, considering a dataset from a GE HDx MRI system with an 8 channel receive only head RF coil (GE Medical Systems, Milwaukee, WI). The acquisition scheme uses a variable density spiral with 377 interleaves using FISP based α [32] and a constant $TR = 10$ms. The excitation sequence length is $L = 1000$. In this experiment, we have FOV $= 22.5 \times 22.5\text{cm}^2$ with a 5mm slice thickness. The EPG model is used for the reconstructions with an inversion time (TI) of 18ms and an Echo Time $TE = 1.902$ms. The scanned objects are the tubes 1, 5 and 9 of the EUROSPIN phantom. We reconstruct the parameter maps with two spatial resolutions: the first one at 180×180 with an undersampling ratio of $N/Q = 44.8753$, and the second one at 40×40 with an undersampling ratio of $N/Q = 20.6869$ to introduce the PV. Note that for the 40×40 reconstruction, only the Fourier samples corresponding to the target resolution are used. Reconstructing for higher spatial resolution would introduce high frequency artefacts as shown in [28]. An acquisition without the tubes is performed to estimate $\sigma_{\mathbf{Y}}$ and compute a lower bound on the iSNR. More precisely, using the triangle inequality, since $\|\mathbf{Y}\|_2 \geq \|\boldsymbol{\eta}\|_2$, we have $\text{iSNR} \geq 20 \log\left((\|\mathbf{Y}\|_2 - \|\boldsymbol{\eta}\|_2)/(\sqrt{QLC}\sigma_{\mathbf{Y}})\right) = 64.73$dB, where $\mathbf{Y}$ corresponds to the measurements with the tubes, $\boldsymbol{\eta}$ corresponds to the measurements without the tubes, and the value $\sigma_{\mathbf{Y}}$ is the standard deviation of $\boldsymbol{\eta}$.

TABLE 4.2: Comparison between the parameters obtained with GAP-MRF and the EUROSPIN phantom values.

	Phantom Values		180×180		40×40	
	T_1	T_2	T_1	T_2	T_1	T_2
Tube 1	200 ± 6	52 ± 1.6	197.0	93.9	195.4	96.0
Tube 5	450 ± 13.5	94 ± 2.8	455.9	159.8	459.5	168.1
Tube 9	754 ± 22.6	116 ± 3.5	766.3	199.2	757.5	199.9

The box in red shows a PV voxel artificially created by reconstructing a lower resolution image. As predicted by the corresponding high resolution maps, this voxel is formed by a linear combination of the Tubes 1 and 9.

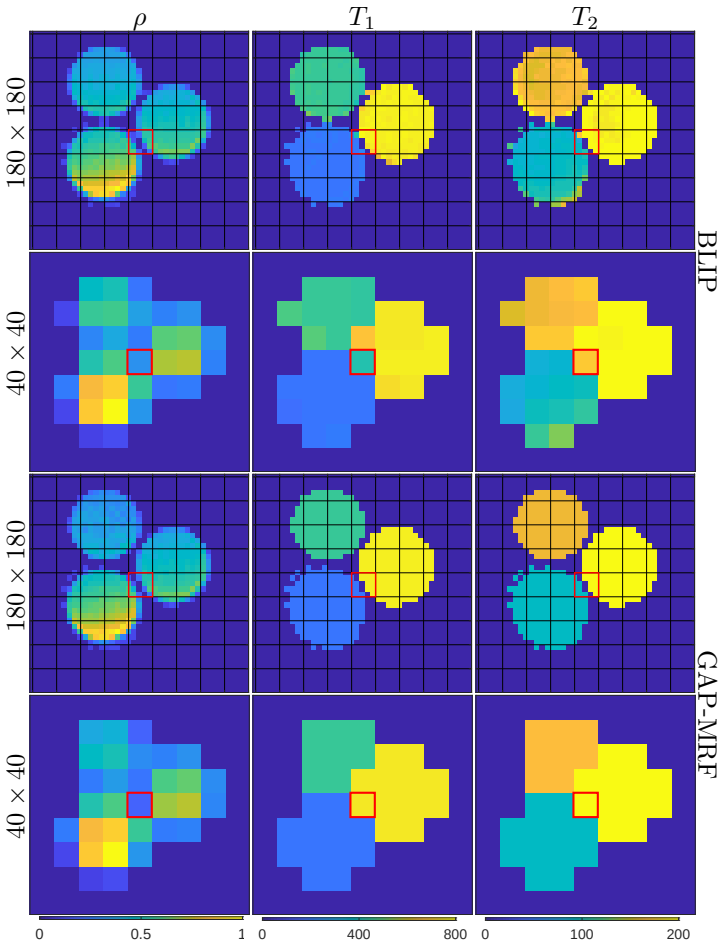


FIGURE 4.9: Dominant tissue parameter maps corresponding to the EUROSPIN phantom dataset. From first to the last column: 180×180 BLIP reconstruction trimmed to 41×41 voxels, 40×40 BLIP reconstruction trimmed to 9×9 voxels, 180×180 GAP-MRF reconstruction trimmed to 41×41 voxels, 40×40 GAP-MRF reconstruction trimmed to 9×9 voxels. From first to the last row: normalised proton density, T_1, and T_2.

In Fig. 4.9, we show a comparison of the reconstructions with two different spatial resolutions. The T_1 and T_2 lower resolution maps of BLIP show a variation introduced by the PVE. A clear example of the PVE is the voxel in the red box where two tissues appear, the BLIP method shows a parameter mismatch. Note that the parameters predicted by BLIP suggest that the voxel contains the same substance as Tube 5, contrary to the true composition (Tubes 1 and 9). GAP-MRF reconstructions do not show this behaviour since we take the PV into account in the model. Note that GAP-MRF is more sensitive to noise as shown in the simulations, this may explain small artefacts in the proton density maps.

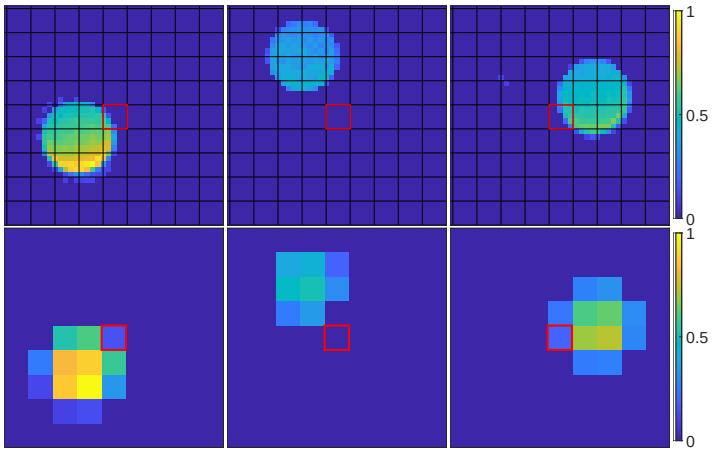


FIGURE 4.10: Normalised proton density maps corresponding to the EUROSPIN phantom dataset. The first and second row correspond to the 180×180 reconstruction trimmed to 41×41, and the 40×40 reconstruction trimmed to 9×9, respectively. The corresponding T_1 and T_2 can be seen in Table 4.2. From left to right the columns correspond to Tube 1, Tube 5 and Tube 9 of the EUROSPIN phantom.

The T_1 values in Table 4.2 are in agreement with the values of the EUROSPIN phantom. The T_2 values are higher than expected. As seen in Fig. 4.9, BLIP results show the same increased T_2, suggesting that the errors may be related to the acquisition parameters specifically to the constant TR as shown in [67].

In Fig. 4.10, the normalised proton density maps reconstructed by GAP-MRF with two different resolutions can be seen. As highlighted by the red box, the PV voxel in the low resolution reconstruction has values different than 0 in the maps corresponding to Tube 1 and 9, this is in agreement with the high resolution maps.

4.5.3.2 *In vivo* brain dataset with spiral sampling

This dataset was acquired by self-experimentation on our team members (all experts in MRI). Since these experiments are not intended to be qualified as a clinical investigation, they do not require any formal IRB approval according to the German Act on Medical Devices (Medical Device Act, MDA). The self-experiments were performed on a device that has already met the requirements of the assessment procedure of conformity, certifying its safety and functionality for the intended purpose (aka "CE

marking of MR scanner"). The experiments were neither invasive nor stressful, therefore, they fully comply with internal GE and German/EU regulations. The scanning for this dataset was performed on a GE HDx MRI system with an 8 channel receive only head RF coil (GE Medical Systems, Milwaukee, WI). The acquisition scheme uses a variable density spiral with 89 interleaves using FISP based α and TR as in [32]. The excitation sequence length is $L = 1000$. In this experiment, we have FOV $= 22.5 \times 22.5\text{cm}^2$ and the spatial resolution is 180×180 voxels, with a 5mm slice thickness. The undersampling ratio is $N/Q = 89.53$. The EPG model is used for the reconstructions with a TI of 18ms and a TE of 2ms. The reconstruction for BLIP and GAP-MRF was accelerated with the SVD compression in the time domain described in [26, 28] using 30 eigenvectors.

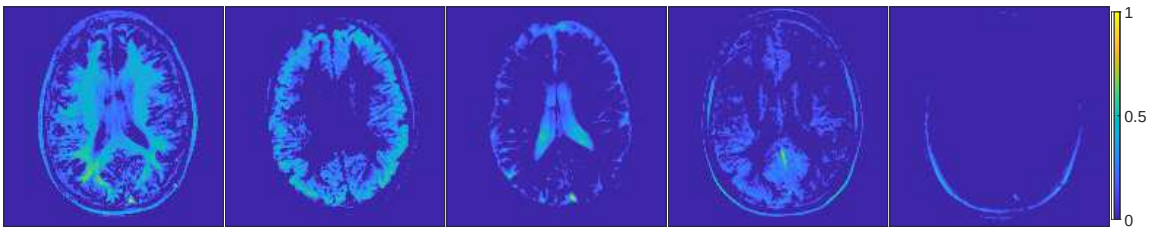


FIGURE 4.11: Normalised proton density maps corresponding to the brain dataset with spiral sampling. The corresponding T_1 and T_2 can be seen in Table 4.3. The reconstructions were trimmed from 180×180 to 151×151 voxels. From left to right and top to bottom, the figures correspond to WM, GM, CSF, muscle and fat.

TABLE 4.3: Comparison between the parameters obtain with GAP-MRF for the brain dataset with spiral sampling and the reported values in [68].

	Values reported in [68]		GAP-MRF	
	T_1	T_2	T_1	T_2
WM	781±61	65±6	758.7	42.1
GM	1193±65	109±11	872.4	67.3
CSF			1658.5	799.8
Muscle	1100±59	44±9	1218.0	23.2
Fat	253±42	68±4	325.5	68.1

In Fig. 4.11, we can observe the resulting proton density maps provided by the GAP-MRF algorithm and the Table 4.3 shows a comparison between the parameters reported in [68] and the parameters obtained by GAP-MRF. The WM, GM and Fat parameters obtained by GAP-MRF slightly differ from those reported for MRF sequences but the values are in agreement with the parameters of other qMRI methods

reported in [68]. The muscle parameters are far from the expected values. This could be due to the small number of pure voxels that are not sufficient to accurately estimate the parameters. We believe that choosing better acquisition parameters $\mathbf{\Gamma}$ to make the elements of the dictionary more distant in the ℓ_2-norm sense can significantly improve the accuracy of the parameters. Also, inaccuracies in the model such as calibration or motion in the acquisition can produce artefacts in the reconstruction. In order to show the importance of the phase compensation in the real data, we present the results without phase compensation. As seen in Fig. 4.12, due to the phase errors there are voxels within the brain without proton density.

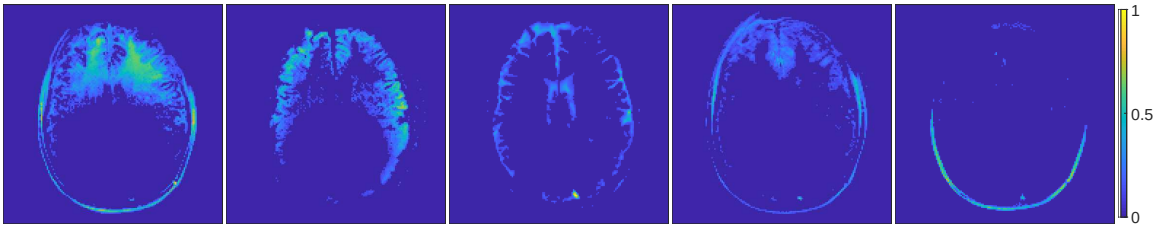


FIGURE 4.12: Normalised proton density maps corresponding to the brain dataset with spiral sampling. The reconstructions were trimmed from 180×180 to 151×151 voxels. From left to right and top to bottom, the figures correspond to WM, GM, CSF, muscle and fat.

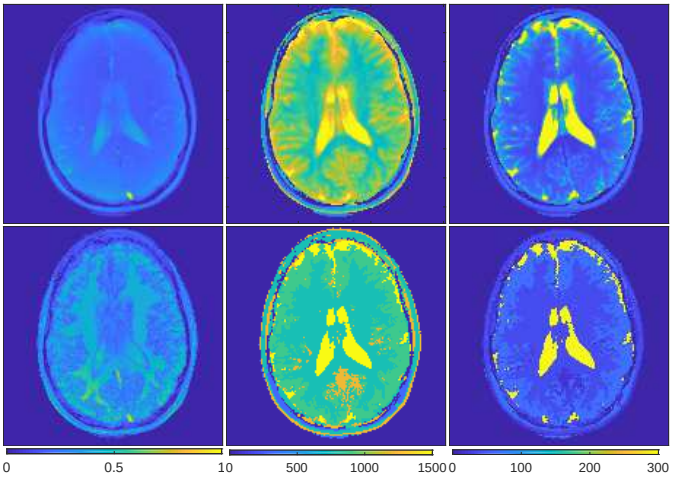


FIGURE 4.13: Dominant tissue parameter maps corresponding to the brain dataset with spiral sampling. The first row corresponds to the BLIP reconstructions and the second row to the GAP-MRF reconstructions. The reconstructions were trimmed from 180×180 to 151×151 voxels. From left to right, the columns correspond to the normalised proton density, T_1 and T_2. The values of T_1 and T_2 are capped to 1500ms and 300ms respectively.

In Fig. 4.13, the T_1 and T_2 maps reconstructed by BLIP show a smooth transition from one tissue to another (similar to the simulated phantom). Moreover, the proton density map reconstructed by BLIP does not provide any information on the tissue distribution. On the contrary, GAP-MRF reconstructions show abrupt transitions in the T_1 and T_2 maps of the dominant tissues. In addition, the proton density map shows more structure than BLIP, but not all the tissue structures are appreciated compared to the normalised proton density maps. Note that the voxels with higher proton density values indicate the pure voxels, and the voxels with reduced values, which are observed at tissue interfaces, indicate the dominant tissue that occupying only a fraction of the voxel.

4.5.3.3 *In vivo* brain dataset with EPI sampling

The scanning for this dataset has been performed on a 3T GE MR750w scanner with a 12 channel receive only head RF coil (GE Medical Systems, Milwaukee, WI). The study was approved by the local ethics committee. The used acquisition scheme was 16-shot EPI-MRF on a healthy volunteer using a variable flip angle α ramp, ranging from 1° to 70°. The excitation sequence length is $L = 500$. The repetition time TR was set to 16ms. In [67], it was shown to be as effective at estimating the MRF parameters but had better sensitivity than the FISP sequence in [32]. The acquisition bandwidth (BW) = 5kHz and the Field of View (FOV) = $22.5 \times 22.5\text{cm}^2$. The spatial resolution is 128×128 voxels, with a 5mm slice thickness. The undersampling ratio is $N/Q = 16$. The EPG model is used for the reconstructions with an Inversion Time (TI) of 18ms and an Echo Time (TE) of 3.5ms. The acquisition time for the slice was 9s. A reference scan with null Gy gradient was performed for phase correction of EPI raw data.

TABLE 4.4: Comparison between the parameters obtain with GAP-MRF for the brain dataset with EPI sampling and the reported values in [68].

	Values reported in [68]		GAP-MRF	
	T_1	T_2	T_1	T_2
WM	781±61	65±6	762.6	67.2
GM	1193±65	109±11	1116.6	107.1
CSF			2391.1	856.2

FIGURE 4.14: Normalised proton density maps corresponding to the brain dataset with EPI sampling. The reconstructions were trimmed from 128×128 to 89×89 voxels. The corresponding T_1 and T_2 can be seen in Table 4.4. From left to right, the figures correspond to WM, GM and CSF.

In Fig. 4.14, we can observe the resulting proton density maps provided by the GAP-MRF algorithm and the Table 4.4 shows a comparison between the parameters reported in [68] for MRF FISP sequences and the parameters obtained by GAP-MRF. CSF values are not reported for the MRF FISP sequence. The WM parameters are similar to the ones reported in [68] and the GM T_1 is slightly lower than the reported one. We believe that the lack of pure voxels (due the spatial resolution) made the approach unable to find the other tissues.

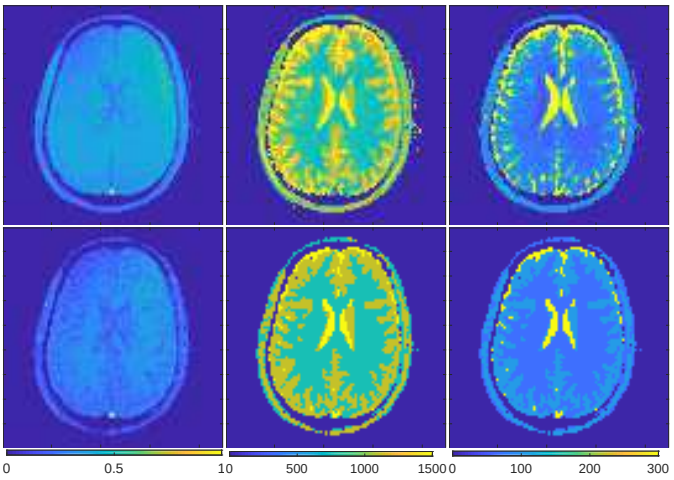


FIGURE 4.15: Dominant tissue parameter maps corresponding to the brain dataset with EPI sampling. The first row corresponds to the BLIP reconstructions and the second row to the GAP-MRF reconstructions. The reconstructions were trimmed from 128×128 to 89×89 voxels. From left to right, the columns correspond to the normalised proton density, T_1 and T_2. The values of T_1 and T_2 are capped to 1500ms and 300ms respectively.

In Fig. 4.15, the T_1 and T_2 maps reconstructed by BLIP shows a smooth transition from one tissue to another. On the contrary, GAP-MRF reconstructions show abrupt transitions in the T_1 and T_2 maps of the dominant tissues.

Chapter 5

Self-Calibration and Imaging

The reconstruction quality of all the computational imaging techniques highly depends on model accuracy. In MRI, coil sensitivity calibration is a crucial step in the reconstruction process to obtain accurate results. Usual MRI self-calibration methods, reconstructing independently the time acquisitions, are not suitable for highly undersampled MRF data. We propose the first self-calibration method for MRF, exploiting the correlation in the time acquisitions, the spatial regularity of the magnetisation images and the smoothness of the coil sensitivity maps. Part of this work was presented in the Biomedical and Astronomical Signal Processing Group (BASP) Frontiers Workshop in 2019.

Fixing the resolution of the reconstructed volume, not only partial volumes are introduced to the model but also coil sensitivity errors. We propose the following self-calibration and imaging problem to best approximate the sensitivity coils and reconstruct the quantitative parameter maps:

$$\underset{\mathbf{M}\in\mathcal{B}_{\mathcal{S}_+}(\mathbf{\Phi}),\mathbf{S}\in\mathbb{C}^{N\times C}}{\text{minimise}} \; F\left(\mathbf{S},\mathbf{M}\right)+r_1 G_{\text{TV}}\left(\mathbf{M}\right)+r_2 G_{\mathbf{S}}\left(\mathbf{S}\right), \tag{5.1}$$

where

$$F(\mathbf{S},\mathbf{M})=1/2\|\mathbf{Y}-h_{\mathbf{S}}(\mathbf{S},\mathbf{M})\|_2^2, \tag{5.2}$$

$h_{\mathbf{S}}\colon \mathbb{C}^{N\times C}\times\mathbb{C}^{N\times L}\to\mathbb{C}^{Q\times L\times C}$ is the linear mapping defining the complete acquisition, i.e. (3.1), $r_1>0$ and $r_2>0$ are parameters used to weight the regularisation terms, G_1 is the function defined in (3.29),

$$G_{\mathbf{S}}\left(\mathbf{S}\right)=\sum_{c=1}^{C}\left\|\mathbf{\Psi}\left(\mathbf{S}_{:,c}\right)\right\|_1, \tag{5.3}$$

where $\mathbf{\Psi}$ is a sparsity basis that should be chosen according to the CS theory. In this manuscript, we propose the two-dimensional Discrete Cosine Transform (DCT). An example of the DCT transform of a coil sensitivity map can be seen in Figure 5.1. Note that the coil sensitivity map is sparse in the DCT domain.

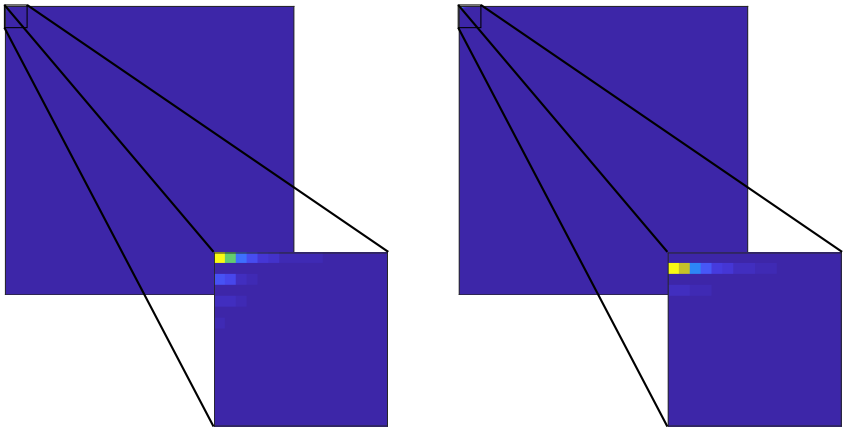


FIGURE 5.1: Example of the DCT of a coil sensitivity map (Left) Absolute value of the real part (Right) Absolute value of the imaginary part.

There are four ingredients in problem (5.2), firstly we have the data fidelity term F that is used to impose data consistency, then we have the non-convex set $\mathcal{B}_{\mathcal{S}_+}$ constraining the signal to be a linear combination of B elements, the regularisation term G_{TV} that promotes sparse solutions of the magnetisation images gradient and finally we have the regularisation term $G_{\mathbf{S}}$ that promotes sparse solutions of the sensitivity coils in the DCT domain.

Algorithm 9 Self-Calibrated GAP-MRF

Input: $\mathbf{S}^{(0)} \in \mathbb{C}^{N\times C}$, $\mathbf{M}^{(0)} \in \mathcal{B}_{\mathcal{S}_+}(\mathbf{\Phi})$, $(I, J) \in \mathbb{N}_*^2$, $(\mu, \nu) \in (\mathbb{R}_*^+)^2$
for $k = 0, 1, \ldots$ **do**
  $\widetilde{\mathbf{M}}^{(0)} = \mathbf{M}^{(k)}$ and $\widetilde{\mathbf{S}}^{(0)} = \mathbf{S}^{(k)}$
  for $i = 0, 1, \ldots, I-1$ **do**
    $\widetilde{\mathbf{M}}^{(i+1)} \approx \text{prox}_{r_1\mu^{(i)}G_{\text{TV}}(\cdot)}\left(\mathcal{P}_{\mathcal{B}_{\mathcal{S}_+}(\mathbf{\Phi})}\left(\widetilde{\mathbf{M}}^{(i)} - \mu^{(i)}\nabla_{\mathbf{M}}F\left(\mathbf{S}^{(k)}, \widetilde{\mathbf{M}}^{(i)}\right)\right)\right)$
  $\mathbf{M}^{(k+1)} = \widetilde{\mathbf{M}}^{(I)}$
  for $j = 0, 1, \ldots, J-1$ **do**
    $\widetilde{\mathbf{S}}^{(j+1)} = \text{prox}_{r_2\nu G_{\mathbf{S}}(\cdot)}\left(\widetilde{\mathbf{S}}^{(j)} - \nu\nabla_{\mathbf{S}}F\left(\widetilde{\mathbf{S}}^{(j)}, \mathbf{M}^{(k+1)}\right)\right)$
  $\mathbf{S}^{(k+1)} = \widetilde{\mathbf{S}}^{(J)}$

To solve (5.1), we use an alternating forward-backward approach [3], described in Algorithm 9, where $\nabla_{\mathbf{M}}F$ (resp. $\nabla_{\mathbf{S}}F$) denotes the partial gradient of F with respect

to $\mathbf{M}$ (resp. $\mathbf{S}$). Note that the projection in the step 5 of Algorithm 9 is performed approximately using the same procedure as GAP-MRF. Under technical conditions [3], the sequence $(\mathbf{M}^{(k)}, \mathbf{S}^{(k)})_{k\in\mathbb{N}}$ generated by Algorithm 9 is guaranteed to converge to a critical point of the objective function of interest. Note that since $F(\mathbf{S}, \mathbf{M}) = F(\mathbf{S}/a, a\mathbf{M})$, where $a \neq 0$, the solution obtained by Algorithm 9 can be a scaled version of the $\mathbf{S}$ and $\mathbf{M}$ (changing the problem depending on the initialisation). To address this problem, at each iteration, we normalise our solutions by the factor $a = \|\mathbf{S}_{z,:}\|_2$ where z is the index of the central voxel in the volume.

5.1 Simulations

Similarly to the previous chapter, we introduce the PVE, we use blocks of 2×2 voxels to form a lower resolution phantom containing PV voxels. For simplicity, we keep the resolution to 128×128 voxels, and we generate the magnetisation sequence. We consider eight coils (i.e. $C = 8$) and the corresponding sensitivity maps can be seen in Fig. 5.2.

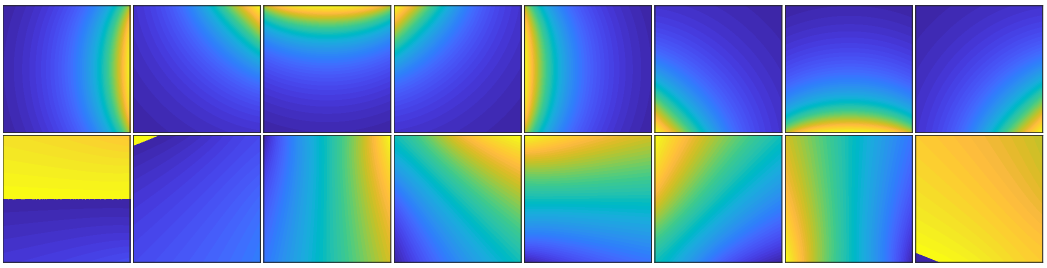


FIGURE 5.2: Coil sensitivity maps used in the simulated phantom experiments. First row corresponds to the magnitude and the second row corresponds to the phase.

The EPG formalism is used for the non-linear mapping, with the flip angles α and repetition times TR as described in [32] and can be seen in Fig. 5.3. Two sampling patterns are considered, an Echo-planar Imaging (EPI) [34, 35] sampling used in [27]. and the spiral sampling used in [2]. In particular, the EPI sampling is performed using equally spaced lines of the k-space and the spiral sampling scheme uses a variable density spiral with 89 interleaves. We compare the reconstructions of GAP-MRF

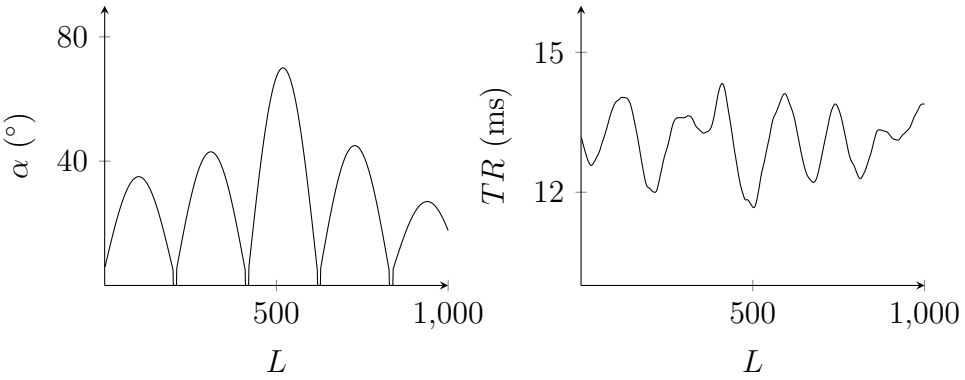


FIGURE 5.3: Acquisition parameters used in the simulated phantom experiments. (Left) Flip angles. (Right) Repetition times.

algorithm and the Self-calibrated GAP-MRF. We initialise the coil sensitivity maps with the calibration method described in [69], using the temporal compressed magnetisation images $h_{\mathbf{S}}^{\dagger}(\mathbf{Y})\widehat{\mathbf{V}}$ with 30 eigenvectors. The obtained proton density maps were normalised in a voxel-wise fashion the represent the tissue proportion.

TABLE 5.1: Parameter values of the reconstructions in Fig. 5.4 corresponding to spiral sampling. The relaxation times are in ms.

	Ground Truth		GAP-MRF		Self-Calibrated GAP-MRF	
	T_1	T_2	T_1	T_2	T_1	T_2
Adipose	530	77	499.6	77.4	504.8	77.5
White Matter	811	77	810.1	76.4	809.9	76.9
Muscle	1425	41	1454.7	41.8	1426.6	41.2
Gray Matter	1545	83	1557.4	82.8	1544.4	82.8
CSF	5012	512	5012.7	512.3	5012.2	510.0

As seen in Figure 5.4-5.5, there is a slight improvement to the PV estimations. It is important to mention that while the tissue proportion maps improvements are small, the relaxation times obtained in Tables 5.1-5.2 by the Self-calibrated GAP-MRF are significantly better.

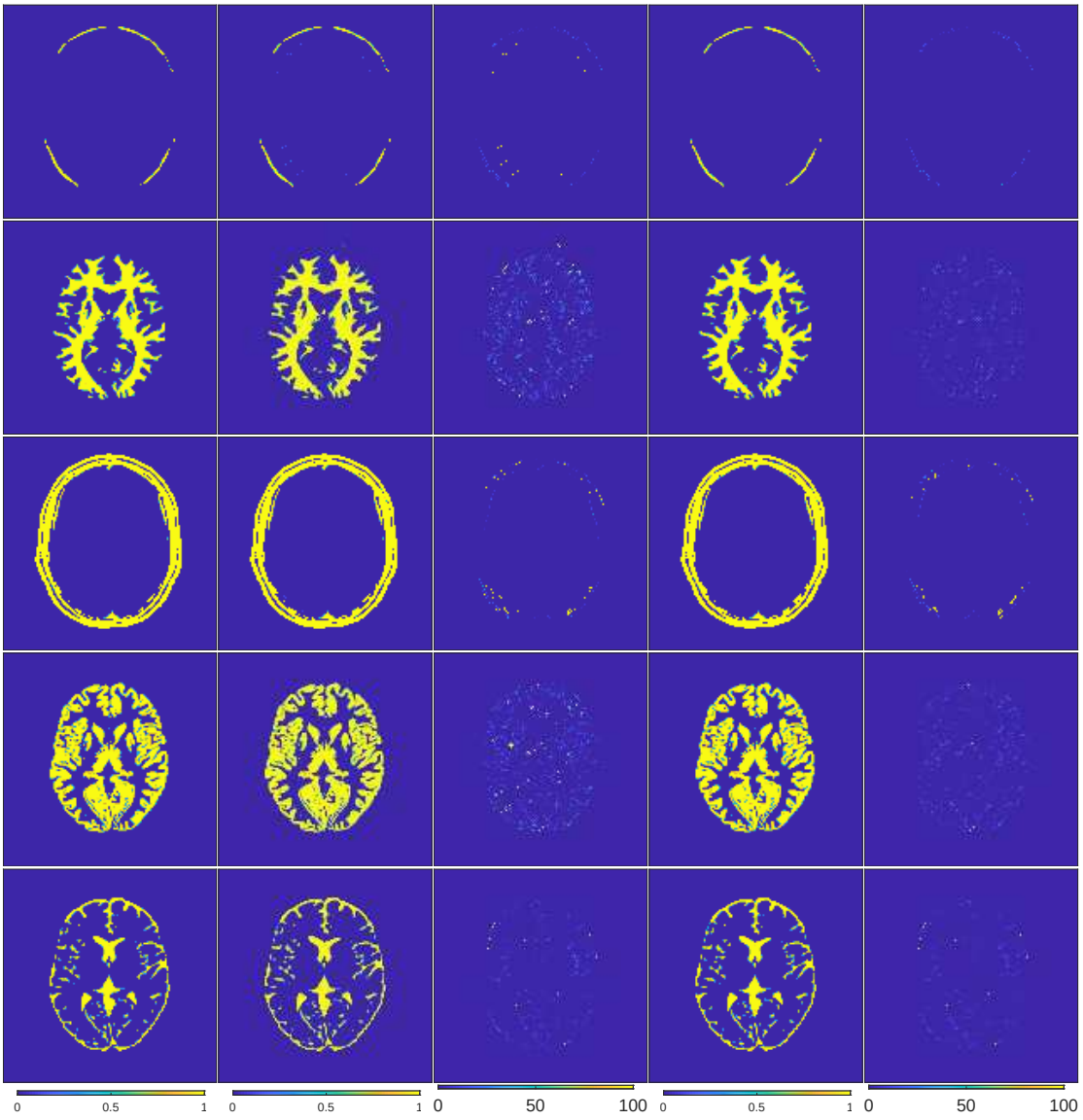


FIGURE 5.4: Tissue proportion maps reconstructed with unknown sensitivity maps with EPI sampling. From first to the last column: Ground truth images, GAP-MRF reconstructions and percentage error between GAP-MRF and the Ground Truth, Self-calibrated GAP-MRF reconstructions and percentage error between Self-Calibrated GAP-MRF and the Ground Truth. From first to the fifth row: Adipose, WM, muscle, GM and CSF. Sixth row: Proton density sum of all other matched elements that are not in the 15% range of the ground truth elements. The corresponding T_1 and T_2 values are given in Table 5.1.

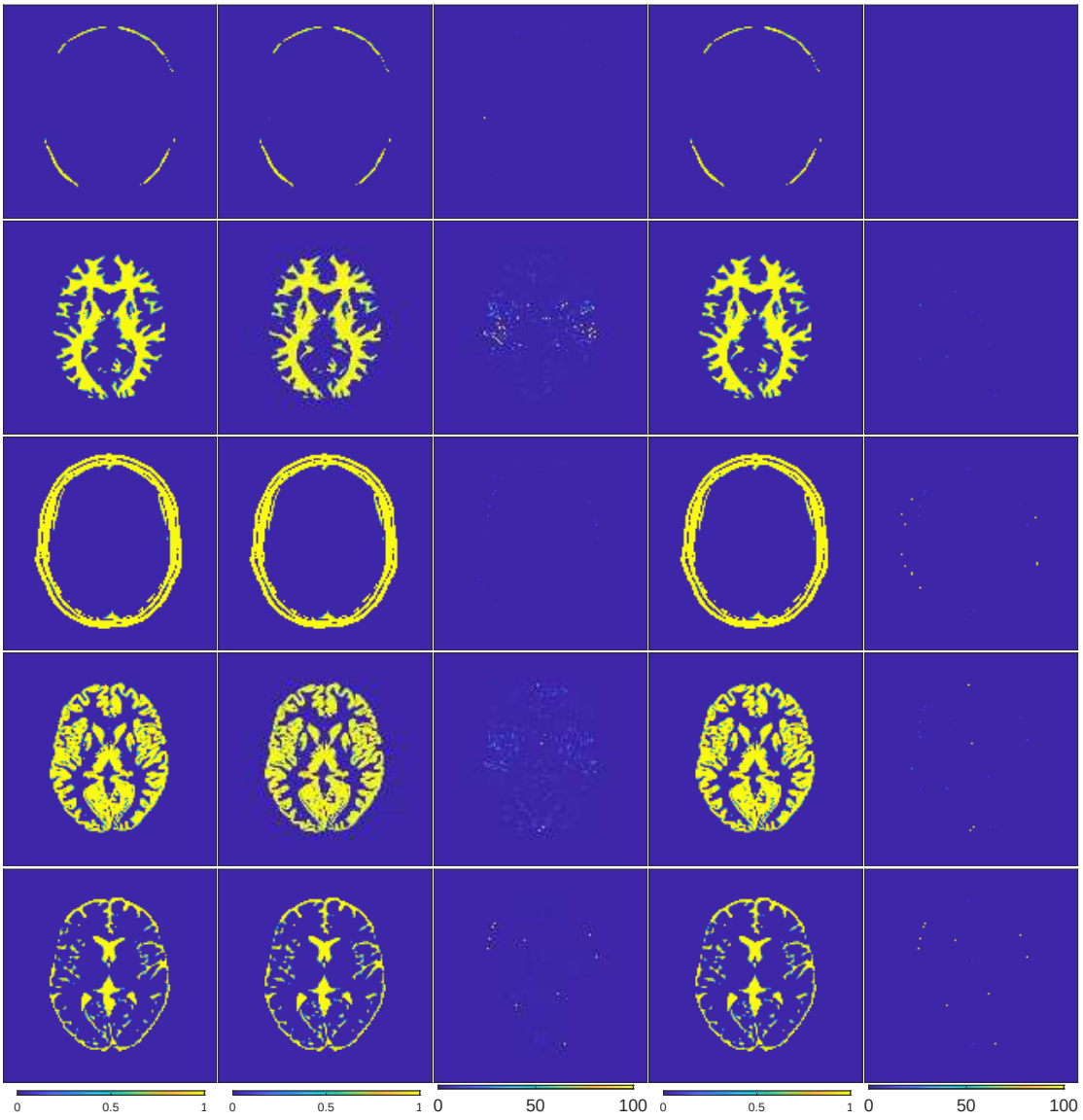


FIGURE 5.5: Tissue proportion maps reconstructed with unknown sensitivity maps with EPI sampling. From first to last column: Ground truth images, GAP-MRF reconstructions and percentage error between GAP-MRF and the Ground Truth, Self calibrated GAP-MRF reconstructions and percentage error between Self-Calibrated GAP-MRF and the Ground Truth. From first to fifth row: Adipose, WM, muscle, GM and CSF. Sixth row: Proton density sum of all other matched elements that are not in the 15% range of the ground truth elements. The corresponding T_1 and T_2 values are given in Table 5.2.

TABLE 5.2: Parameter values of the reconstructions in Fig. 5.5 corresponding to EPI sampling. The relaxation times are in ms.

	Ground Truth		GAP-MRF		Self-Calibrated GAP-MRF	
	T_1	T_2	T_1	T_2	T_1	T_2
Adipose	530	77	523.7	74.9	530.8	76.8
White Matter	811	77	802.0	75.6	810.5	76.9
Muscle	1425	41	1402.4	40.3	1426.2	41.1
Gray Matter	1545	83	1527.1	81.6	1544.4	83.0
CSF	5012	512	5001.7	510.3	5011.8	512.0

5.2 Realistic simulations

While the simulations in the previous section were used as proof of concept to evaluate Self-calibrated GAP-MRF, they are not realistic. In order to produce a more realistic scenario, we have generated the measurements using the high-resolved phantom [33] (256x256 voxels) without partial volume and the corresponding sensitivity coils. The measurements were taken only from the centre of the k-space corresponding to an image with a quarter of the voxels (128x128 voxels). The sampling pattern of the first acquisition instance $L = 1$ can be seen in Figure 5.6. Note that we are not only introducing PV but also coil sensitivity inaccuracies. The results of these simulations are presented in Figures 5.7-5.8. Note that we use the ground truth from the simulated phantom in the previous section as reference since for these simulations there is no ground truth.

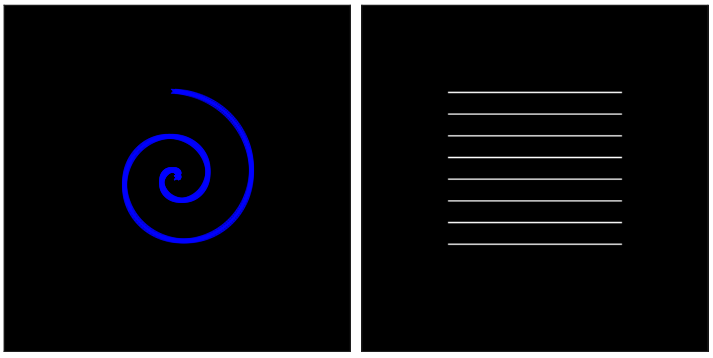


FIGURE 5.6: Sampling patterns for $L = 1$. (Left) Spiral sampling. (Right) EPI sampling.

As seen in Figure 5.7-5.8, there is a slight improvement to the PV estimations. It is important to mention that while the tissue proportion maps improvements are small,

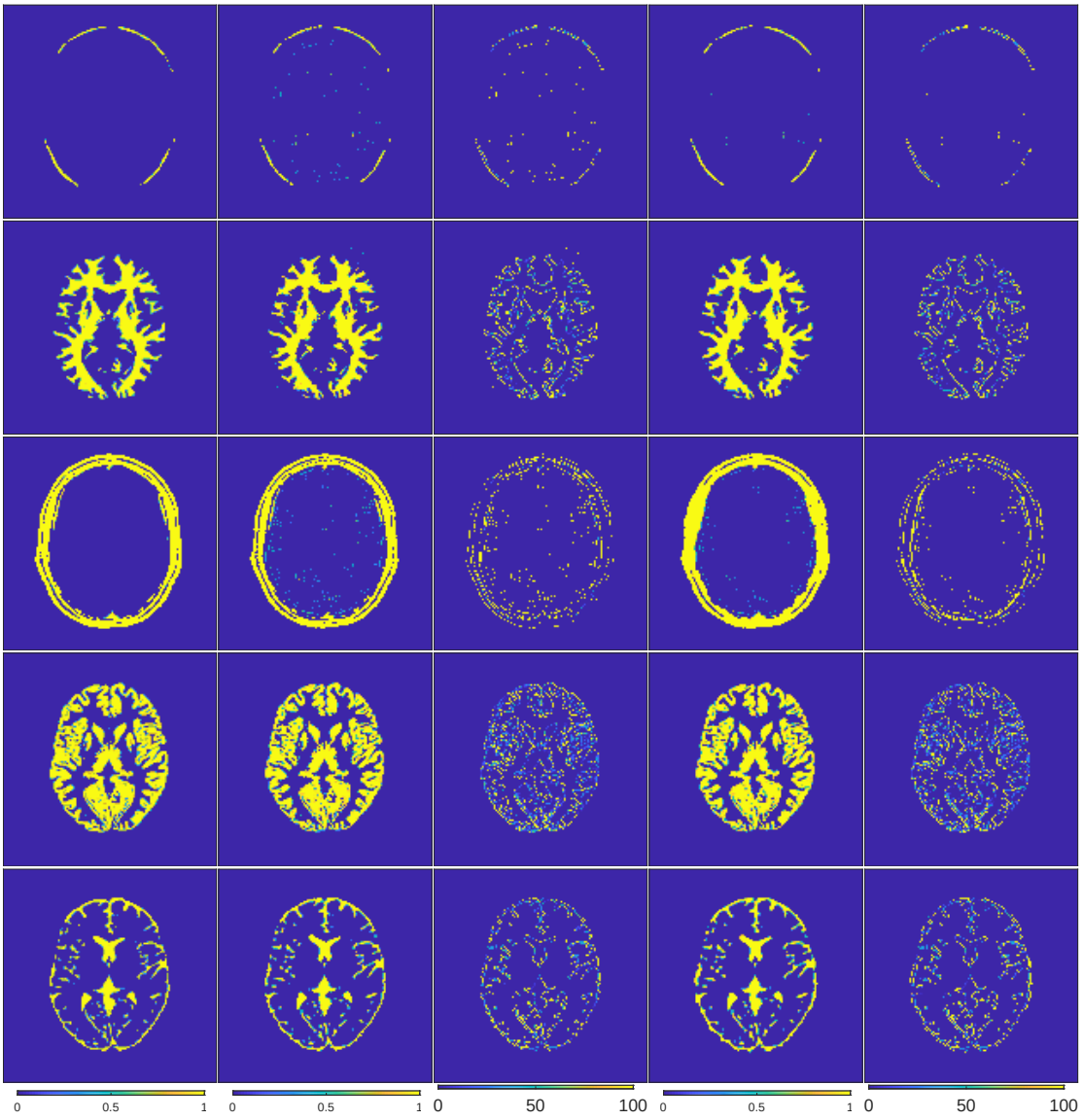


FIGURE 5.7: Tissue proportion maps reconstructed with unknown sensitivity maps with EPI sampling. From first to the last column: Ground truth images, GAP-MRF reconstructions and percentage error between GAP-MRF and the Ground Truth, Self calibrated GAP-MRF reconstructions and percentage error between Self-Calibrated GAP-MRF and the Ground Truth. From first to the fifth row: Adipose, WM, muscle, GM and CSF. Sixth row: Proton density sum of all other matched elements that are not in the 15% range of the ground truth elements. The corresponding T_1 and T_2 values are given in Table 5.3.

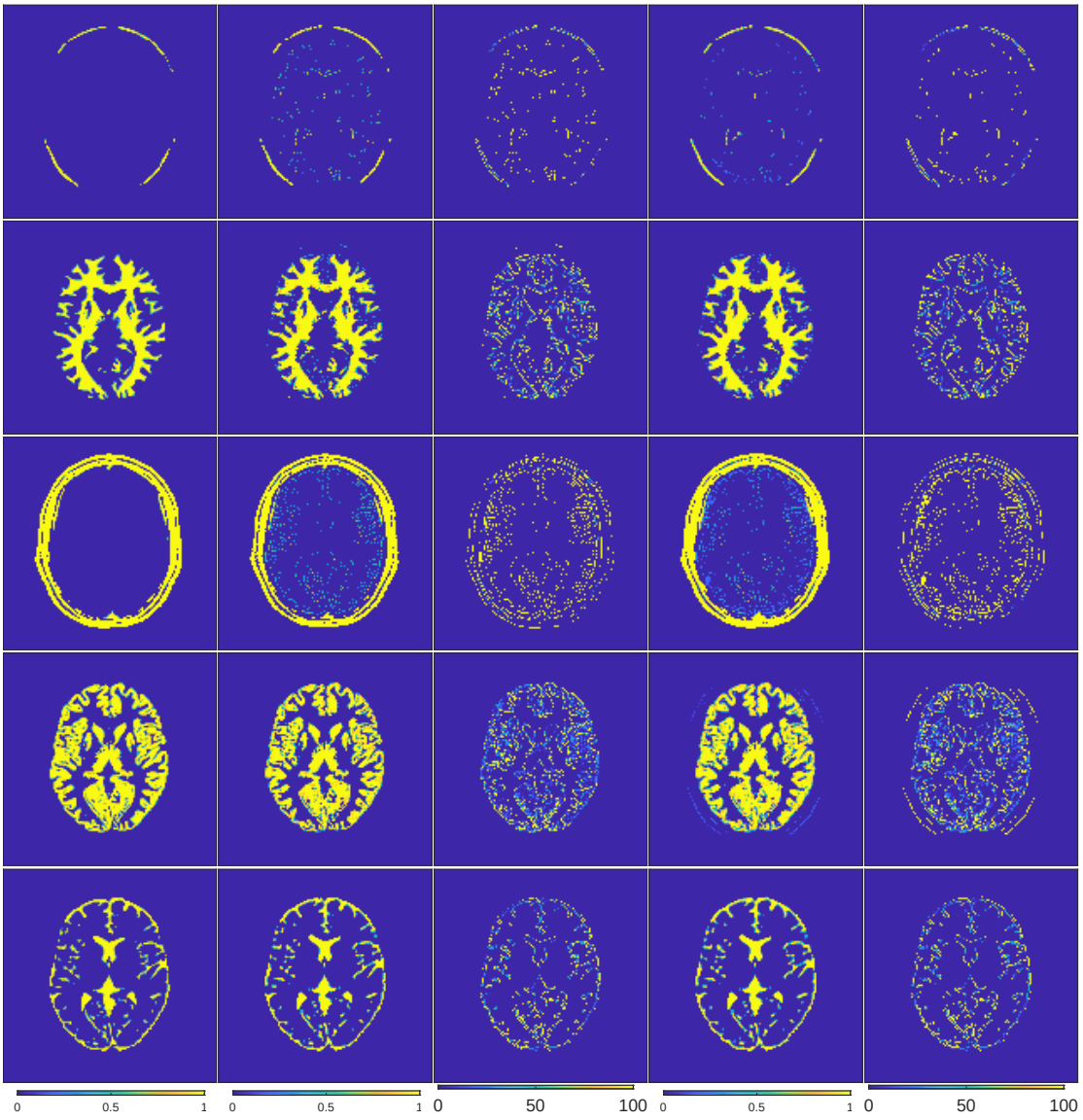


FIGURE 5.8: Tissue proportion maps reconstructed with unknown sensitivity maps with EPI sampling. From first to the last column: Ground truth images, GAP-MRF reconstructions and percentage error between GAP-MRF and the Ground Truth, Self calibrated GAP-MRF reconstructions and percentage error between Self-Calibrated GAP-MRF and the Ground Truth. From first to the fifth row: Adipose, WM, muscle, GM and CSF. Sixth row: Proton density sum of all other matched elements that are not in the 15% range of the ground truth elements. The corresponding T_1 and T_2 values are given in Table 5.4.

TABLE 5.3: Parameter values of the reconstructions in Fig. 5.7 corresponding to spiral sampling. The relaxation times are in ms.

	Ground Truth		GAP-MRF		Self-Calibrated GAP-MRF	
	T_1	T_2	T_1	T_2	T_1	T_2
Adipose	530	77	449.6	77.7	447.5	79.6
White Matter	811	77	806.9	77.2	814.6	77.4
Muscle	1425	41	1422.3	42.1	1495.3	41.2
Gray Matter	1545	83	1486.4	80.6	1536.8	82.7
CSF	5012	512	5093.6	532.9	4996.9	510.0

TABLE 5.4: Parameter values of the reconstructions in Fig. 5.8 corresponding to EPI sampling. The relaxation times are in ms.

	Ground Truth		GAP-MRF		Self-Calibrated GAP-MRF	
	T_1	T_2	T_1	T_2	T_1	T_2
Adipose	530	77	496.6	80.2	520.0	77.4
White Matter	811	77	819.5	80.4	803.0	77.2
Muscle	1425	41	1382.9	42.7	1406.8	46.0
Gray Matter	1545	83	1588.2	87.4	1593.5	84.9
CSF	5012	512	5014.2	523.1	5065.8	513.5

the relaxation times obtained in Tables 5.3-5.4 by the Self-calibrated GAP-MRF are closer to the real values. Note that by not having a high resolved coil sensitivity maps both methods detect mixture of tissues in regions pure tissues. This suggest that both methods are highly sensitive to model inaccuracies.

Chapter 6

Conclusions and Future Work

In Chapter 3, we have presented novel algorithms for MRF based on the framework in [27] that allow accelerated and super-resolution reconstructions. In particular, these algorithm handles the projection onto $\mathcal{B}^+$ efficiently by updating the dictionary at each iteration. The super-resolution reconstructions rely on the spatial regularisation term added to the minimisation problem. The algorithms developed in Chapter 3 (i.e. ABLIP, ℓ_1-rABLIP and TV-ABLIP) can be explained in the same framework as GAP-MRF algorithm in Chapter 4.

Based on the framework in [27], we developed our partial volume model in Chapter 4. Our algorithm provides a way to explore the manifold of magnetic resonance fingerprints without densely sampling $\mathcal{M}$. For this reason, the algorithm is memory efficient and the algorithmic structure allows parallel implementations. The proposed model assumes that the number of independent tissues in the imaged volume is upper bounded, and that each tissue has a minimum number of pure voxels. Also, the parameters of each tissue should be sufficiently different to be distinguished. Finally, we assume that the combination of the sampling patterns should cover most of the k-space to avoid high frequency artefacts or the minimisation function should incorporate a spatial regularisation term as proposed in Chapter 5.

The simulation results presented in this manuscript show that the proposed GAP-MRF method can achieve accurate reconstructions with very short pulse sequences in the low input noise scenario. It also performs well when the iSNR is greater than 30dB. We also present results obtained with *in vivo* datasets. Some parameters differ slightly to the reported in the literature, but the structure seen in the proton densities maps suggests that this approach can provide additional information that can be useful for diagnosis.

As an extension of GAP-MRF, we have proposed a Self-Calibration and Imaging algorithm in Chapter 5. The target minimisation problem exploits the correlation in the time acquisitions, the spatial regularity of the magnetisation images and the smoothness of the coil sensitivity maps. The reconstruction algorithm is based on the block coordinate forward-backward algorithm in [3]. We show that self-calibration improves the accuracy of the quantitative parameters (i.e. relaxation times). We acknowledge that the Self-Calibrated GAP-MRF is highly sensitive to model inaccuracies. As mentioned in Chapter 5, by fixing the resolution not only partial volumes are introduced but also coil sensitivity errors.

The next step is to evaluate the PV reconstructions with a real PV phantom in the scanner and a full brain reconstruction to provide enough pure voxels to accurately estimate the true parameters. In particular, an interesting point would be to evaluate the behaviour of GAP-MRF in presence of a pathology. A pathology can be seen as a distinct additional tissue. Therefore, since the number of tissues is estimated along the iterations, if the pathology is represented by enough pure voxels, it should be detected by the algorithm exactly in the same way as for the other tissues. In addition, we will explore different priors to make the Self-Calibration and Imaging more robust to the model inaccuracies. We will also explore the use of data pre-processing algorithms such as low-pass filtering to mitigate the model inaccuracies. Furthermore, we will evaluate our algorithms with 3-dimensional MRF measurements. Finally, we acknowledge that deep learning is an interesting research direction to accelerate the projection onto $\mathcal{B}_{\mathcal{S}_+}$.

Bibliography

[1] J. Beutel, R. Van Metter, and H. Kundel, *Handbook of Medical Imaging: Physics and Psychophysics*. Handbook of Medical Imaging: Physics and Psychophysics, Society of Photo Optical, 2000.

[2] D. Ma, V. Gulani, N. Seibelich, K. Liu, J. L. Duerk, and M. A. Griswold, "Magnetic resonance fingerprinting," *Nature*, vol. 465, pp. 187–192, 2013.

[3] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, "A block coordinate variable metric forward-backward algorithm," *Journal of Global Optimization*, pp. 1–29, Feb. 2016.

[4] H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, 2017.

[5] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*, vol. 317. Springer Science & Business Media, 2009.

[6] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289–1306, Apr. 2006.

[7] G. Wright, "Magnetic resonance imaging," *IEEE Signal Process. Mag.*, vol. 14, pp. 56–66, 1997.

[8] Z. Liang, P. Lauterbur, I. E. in Medicine, and B. Society, *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*. IEEE Press series in biomedical engineering, SPIE Optical Engineering Press, 2000.

[9] M. Bernstein, K. King, and X. Zhou, *Handbook of MRI Pulse Sequences*. Elsevier Science, 2004.

[10] J. Pauly, P. Le Roux, D. Nishimura, and A. Macovski, "Parameter relations for the Shinnar-Le Roux selective excitation pulse design algorithm (NMR imaging)," *IEEE Transactions on Medical Imaging*, vol. 10, pp. 53–65, March 1991.

[11] J. A. Fessler and B. P. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Transactions on Signal Processing*, vol. 51, pp. 560–574, Feb 2003.

[12] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, pp. 489–509, Feb. 2006.

[13] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.

[14] P. L. Combettes and J.-C. Pesquet, *Proximal Splitting Methods in Signal Processing*, pp. 185–212. New York, NY: Springer New York, 2011.

[15] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[16] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1989.

[17] J. Dunn, "Convexity, monotonicity, and gradient processes in hilbert space," *Journal of Mathematical Analysis and Applications*, vol. 53, no. 1, pp. 145 – 158, 1976.

[18] P. Combettes and V. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1168–1200, 2005.

[19] H. Attouch, J. Bolte, and B. F. Svaiter, "Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods," *Mathematical Programming*, vol. 137, no. 1-2, pp. 91–129, 2013.

[20] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear Gauss–Seidel method under convex constraints," *Operations research letters*, vol. 26, no. 3, pp. 127–136, 2000.

[21] M. Doneva, P. Börnert, H. Eggers, C. Stehning, J. Sénégas, and A. Mertins, "Compressed sensing reconstruction for magnetic resonance parameter mapping," *Magnetic Resonance in Medicine*, vol. 64, no. 4, pp. 1114–1120, 2010.

[22] M. Doneva and M. A. Mertins, "Compressed Sensing in Quantitative MRI," in *MRI*, pp. 51–71, CRC Press, 2018.

[23] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[24] B. Zhao, J. P. Haldar, K. Setsompop, and L. L. Wald, "Optimal experiment design for magnetic resonance fingerprinting," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 453–456, Aug 2016.

[25] M. Weigel, "Extended phase graphs: dephasing, RF pulses, and echoes-pure and simple," *Journal of Magnetic Resonance Imaging*, vol. 41, no. 2, pp. 266–295, 2015.

[26] D. F. McGivney, E. Pierre, D. Ma, Y. Jiang, H. Saybasili, V. Gulani, and M. A. Griswold, "SVD compression for magnetic resonance fingerprinting in the time domain," *IEEE transactions on medical imaging*, vol. 33, no. 12, pp. 2311–2322, 2014.

[27] M. Davies, G. Puy, P. Vandergheynst, and Y. Wiaux, "A Compressed Sensing Framework for Magnetic Resonance Fingerprinting," *SIAM J. Imaging Sci.*, vol. 7, p. 2623, Nov. 2014.

[28] C. C. Cline, X. Chen, B. Mailhe, Q. Wang, J. Pfeuffer, M. Nittka, M. A. Griswold, P. Speier, and M. S. Nadar, "Air-mrf: Accelerated iterative reconstruction for magnetic resonance fingerprinting," *Magnetic Resonance Imaging*, vol. 41, pp. 29 – 40, 2017.

[29] S. F. Cauley, K. Setsompop, D. Ma, Y. Jiang, H. Ye, E. Adalsteinsson, M. A. Griswold, and L. L. Wald, "Fast group matching for mr fingerprinting reconstruction," *Magnetic Resonance in Medicine*, vol. 74, no. 2, pp. 523–528, 2015.

[30] M. Golbabaee, Z. Chen, Y. Wiaux, and M. Davies, "Coverblip: scalable iterative matched-filtering for mr fingerprint recovery," in *Joint Annual Meeting ISMRM-ESMRMB 2018*, 2 2018.

[31] M. Poustchi-Amin, S. A. Mirowitz, J. J. Brown, R. C. McKinstry, and T. Li, "Principles and applications of echo-planar imaging: A review for the general radiologist," *RadioGraphics*, vol. 21, no. 3, pp. 767–779, 2001. PMID: 11353123.

[32] Y. Jiang, D. Ma, N. Seiberlich, V. Gulani, and M. A. Griswold, "MR fingerprinting using fast imaging with steady state precession (FISP) with spiral readout," *Magnetic Resonance in Medicine*, vol. 74, no. 6, pp. 1621–1631, 2015.

[33] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imag.*, vol. 17, pp. 463–468, Jun. 1998.

[34] B. Rieger, F. Zimmer, J. Zapp, S. Weingärtner, and L. R. Schad, "Magnetic resonance fingerprinting using echo-planar imaging: Joint quantification of T1 and T2* relaxation times," *Magnetic Resonance in Medicine*, vol. 78, no. 5, pp. 1724–1733, 2017.

[35] B. Rieger, M. Akçakaya, J. C. Pariente, S. Llufriu, E. Martinez-Heras, S. Weingärtner, and L. R. Schad, "Time efficient whole-brain coverage with MR Fingerprinting using slice-interleaved echo-planar-imaging," *Scientific Reports*, vol. 8, no. 1, p. 6667, 2018.

[36] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2845–2862, Nov. 2001.

[37] E. Y. Pierre, D. Ma, Y. Chen, C. Badve, and M. A. Griswold, "Multiscale reconstruction for mr fingerprinting," *Magnetic Resonance in Medicine*, vol. 75, no. 6, pp. 2481–2492, 2016.

[38] B. Zhao, K. Setsompop, E. Adalsteinsson, B. Gagoski, H. Ye, D. Ma, Y. Jiang, P. Ellen Grant, M. A. Griswold, and L. L. Wald, "Improved magnetic resonance fingerprinting reconstruction with low-rank and subspace modeling," *Magnetic Resonance in Medicine*, vol. 79, no. 2, pp. 933–942, 2018.

[39] M. Doneva, T. Amthor, P. Koken, K. Sommer, and P. Börnert, "Matrix completion-based reconstruction for undersampled magnetic resonance fingerprinting data," *Magnetic Resonance Imaging*, vol. 41, pp. 41 – 52, 2017.

[40] G. Mazor, L. Weizman, A. Tal, and Y. C. Eldar, "Low-rank magnetic resonance fingerprinting," *Medical Physics*, vol. 45, no. 9, pp. 4066–4084, 2018.

[41] J. Tohka, "Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: A review," *World J Radiol*, vol. 6, pp. 855–864, Nov. 2014.

[42] A. Deshmane, D. F. McGivney, D. Ma, Y. Jiang, C. Badve, V. Gulani, N. Seiberlich, and M. A. Griswold, "Partial volume mapping using magnetic resonance fingerprinting," *NMR in Biomedicine*, p. e4082, 2019.

[43] D. McGivney, A. Deshmane, Y. Jiang, D. Ma, C. Badve, A. Sloan, V. Gulani, and M. Griswold, "Bayesian estimation of multicomponent relaxation parameters in magnetic resonance fingerprinting," *Magnetic resonance in medicine*, vol. 80, no. 1, pp. 159–170, 2018.

[44] D. P. Bertsekas, *Nonlinear programming*. Athena scientific Belmont, 1999.

[45] P. Tseng, "Convergence of a Block Coordinate Descent Method for Nondifferentiable Minimization," *Journal of Optimization Theory and Applications*, vol. 109, pp. 475–494, Jun. 2001.

[46] J. Bolte, S. Sabach, and M. Teboulle, "Proximal Alternating Linearized Minimization for Nonconvex and Nonsmooth Problems," *Math. Program.*, vol. 146, pp. 459–494, Aug. 2014.

[47] P. Frankel, G. Garrigos, and J. Peypouquet, "Splitting Methods with Variable Metric for Kurdyka–Łojasiewicz Functions and General Convergence Rates,"

Journal of Optimization Theory and Applications, vol. 165, pp. 874–900, Jun. 2015.

[48] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 43, pp. 129–159, 2001.

[49] E. J. Candes, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59 – 73, 2011.

[50] S. Tang, C. Fernandez-Granda, S. Lannuzel, B. Bernstein, R. Lattanzi, M. Cloos, F. Knoll, and J. Assländer, "Multicompartment magnetic resonance fingerprinting," *Inverse problems*, vol. 34, no. 9, p. 094005, 2018.

[51] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *IEEE Trans. Pattern Anal. Mach. Int.*, vol. 24, pp. 881–892, Jul. 2002.

[52] D. Wipf and S. Nagarajan, "Iterative Reweighted ℓ_1 and ℓ_2 Methods for Finding Sparse Solutions," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 317–329, 2010.

[53] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier analysis and applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[54] T. Blumensath, "Sampling and Reconstructing Signals From a Union of Linear Subspaces," *IEEE Trans. Inform. Theory*, vol. 57, pp. 4660–4671, Jul. 2011.

[55] T. Blumensath and M. E. Davies, "Normalized Iterative Hard Thresholding: Guaranteed Stability and Performance," *IEEE J. Selected Topics Signal Process.*, vol. 4, pp. 298–309, Apr. 2010.

[56] M. Golbabaee, D. Chen, P. A. Gómez, M. I. Menzel, and M. E. Davies, "A deep learning approach for magnetic resonance fingerprinting," *CoRR*, vol. abs/1809.01749, 2018.

[57] E. Hoppe, G. Körzdörfer, T. Würfl, J. Wetzl, F. Lugauer, J. Pfeuffer, and A. K. Maier, "Deep Learning for Magnetic Resonance Fingerprinting: A New Approach for Predicting Quantitative Parameter Values from Time Series," in *GMDS*, pp. 202–206, 2017.

[58] O. Cohen, B. Zhu, and M. S. Rosen, "MR fingerprinting deep reconstruction network (DRONE)," *Magnetic resonance in medicine*, vol. 80, no. 3, pp. 885–894, 2018.

[59] P. Virtue, X. Y. Stella, and M. Lustig, "Better than real: Complex-valued neural nets for MRI fingerprinting," in *2017 IEEE International Conference on Image Processing (ICIP)*, pp. 3953–3957, IEEE, 2017.

[60] G. H.-G. Chen and R. T. Rockafellar, "Convergence Rates in Forward–Backward Splitting," *SIAM J. Optim.*, vol. 7, no. 2, pp. 421–444, 1997.

[61] P. Tseng, "A modified forward-backward splitting method for maximal monotone mappings," *SIAM J. Control Optim.*, vol. 38, no. 2, pp. 431–446, 2000.

[62] C. L. Lawson and R. J. Hanson, *Solving least squares problems*, vol. 15 of *Classics in Applied Mathematics*. Philadelphia, PA: SIAM, 1995.

[63] D. P. Bertsekas, *Nonlinear programming*. Athena scientific Belmont, 1999.

[64] J. Canny, "A Computational Approach to Edge Detection," *IEEE Trans. Pattern Anal. Mach. Int.*, vol. PAMI-8, pp. 679–698, Nov. 1986.

[65] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inform. Theory*, vol. 28, pp. 129–137, Mar. 1982.

[66] A. Benjamin, P. Gómez, M. Golbabaee, Z. Mahbub, T. Sprenger, M. Menzel, M. Davies, and I. Marshall, "Balanced multi-shot EPI for accelerated Cartesian MR Fingerprinting: An alternative to spiral MR Fingerprinting," in *Proc Intl Soc Mag Reson Med*, p. 4265, Feb. 2018.

[67] P. A. Gómez, G. Bounincontri, M. Molina-Romero, J. I. Sperl, M. I. Menzel, and B. H. Menze, "Accelerated parameter mapping with compressed sensing: an alternative to MR Fingerprinting," in *Proc Intl. Soc. Mag. Reson. Med*, 2017.

[68] J. Z. Bojorquez, S. Bricq, C. Acquitter, F. Brunotte, P. M. Walker, and A. Lalande, "What are normal relaxation times of tissues at 3T?," *Magnetic Resonance Imaging*, vol. 35, pp. 69 – 80, 2017.

[69] M. Uecker, P. Lai, M. J. Murphy, P. Virtue, M. Elad, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "ESPIRiT—an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA," *Magnetic resonance in medicine*, vol. 71, no. 3, pp. 990–1001, 2014.